**The Dissertation Committee for Anthony James Romo Certifies that this is the approved version of the following dissertation:**

# Discovery of the Amipurimycin and Miharamycins Biosynthetic Gene Clusters and Insight into the Biosynthesis of Nogalamycin

**Committee:**

Hung-wen (Ben) Liu, Supervisor

Walter L. Fast

Adrian T. Keatinge-Clay

Sean M. Kerwin

Christian P. Whitman

# Discovery of the Amipurimycin and Miharamycins Biosynthetic Gene Clusters and Insight into the Biosynthesis of Nogalamycin

**by**

**Anthony James Romo, Pharm.D.**

**Dissertation**

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

**Doctor of Philosophy**

**The University of Texas at Austin**

**May 2016**

## Dedication

I would like to dedicate this work to my amazingly talented wife, Dr. Dawn Kim-Romo, and our beautiful son, Julian Romo. Thank you for encouraging me to continue when I wanted to quit and for giving me the inspiration to reach farther than my grasp.

"O sol che sani ogne vista turbata,
tu mi contenti sì quando tu solvi,
che, non men che saver, dubbiar m'aggrata."
(Dante, *Inferno*, 11.91-93)

## Acknowledgements

First of all, I want to thank God for my wife and son and for the strength to get through my graduate studies, for, though it is often said at this point in a graduate student's life, there were times when I really could never imagine finishing graduate school!

I would be completely and unforgivably remiss if I did not thank the people who have helped shape my experience. I thank Dr. Zhiwen (Jonathan) Zhang for giving me my first experience in medicinal chemistry research and Dr. Gabby Moreno-Thibodeaux for being my first mentor. Next, I would like to thank Dr. Hung-wen (Ben) Liu for allowing me to work in and ultimately study in his lab, and Dr. Yung-nan Liu for greatly facilitating my progress. Not only was Dr. Liu a knowledgeable and thoughtful advisor, but my time in his lab allowed me to interact with a vibrant and enthusiastic group of graduate students and postdocs, some of whom have made a lifelong impression on me.

I would like to specifically thank Dr. Yasushi Ogasawara for giving me the set of skills that were instrumental to my work; Aoshu Zhong for being an apt protégé and skillful teacher; Geng-Min Lin for his patient guidance and encouragement, as well as for continuing my work; Dr. Chia-I Lin for her advice and expertise which often helped me to think more deeply about my results; and finally, Dr. Mark Ruszczycky for telling me to stop "thinking like a first year graduate student" and for being one of the most thorough editors I have ever met. I will never forget my time in the Liu Lab, and I am glad to join the brilliant alumni who have studied with Dr. Liu.

Finally, I would like to thank my committee members for their patience and commitment to this endeavor.

# Discovery of the Amipurimycin and Miharamycins Biosynthetic Gene Clusters and Insight into the Biosynthesis of Nogalamycin

Anthony James Romo, Ph.D.

The University of Texas at Austin, 2016

Supervisor: Hung-wen (Ben) Liu

Advancements in our ability to obtain high quality bacterial whole genome sequences have increased the rate natural product biosynthetic gene clusters are identified, but these projects require computational methods for gene annotation which are limited by their reliance on comparison to previously annotated genes. Thus, even in a well-defined cluster with a known product it may be difficult to predict how structural characteristics arise if they have no known precedent. Without further work, such limitations will continue to impede the utility of sequenced bacterial genomes. Biosyntheses of the peptidyl nucleoside antibiotics and atypical anthracyclines like nogalamycin are topics which exemplify these challenges.

The peptidyl nucleoside antibiotics (PNAs) are a structurally complex group of natural products with diverse biological activities which could be useful for the development of novel antimicrobials. Amipurimycin and the miharamycins are remarkably similar PNAs elaborated by the bacteria *Streptomyces miharaensis* ATCC 19440 and *Streptomyces novoguineensis* CBS 199.78, respectively. Their dissimilarity to other well-characterized groups of antibiotics has presented a challenge to the study of their unique structural features. Herein, we describe the identification of the amipurimycin and miharamycins biosynthetic gene clusters though a comparative genomics approach.

Besides providing insight into the biosynthesis of the rare amino acids adorning these PNA, our analysis revealed a plausible biosynthetic route to the unique 2-aminopurine nucleobase and suggests the core saccharides are generated by a polyketide synthase.

The anthracyclines are a mainstay of chemotherapy, but their use is limited by fatal cardiotoxicity and tumor resistance. The structures of the anthracyclines are sensitive to modification, as even small changes can ablate their biological activity. Nevertheless, the diversification of anthracyclines through semi- or total syntheses is an ongoing effort. Nogalamycin is rare amongst the anthracyclines because of its extended ring system and unusual glycosylation pattern. It is hoped an understanding of the biosynthesis of nogalamycin could allow the incorporation of its uncommon structural features into other anthracyclines to develop novel compounds with improved actitivty. Towards this end, we investigated the biosynthesis of the amino sugar found in nogalamycin, nogalamine, to clarify ambiguous steps in the reported biosynthetic pathway for this sugar.

# Table of Contents

# List of Tables

# List of Figures

# Chapter 1: Biosynthesis of The Peptidyl Nucleoside Antibiotics and Glycosylation of Anthracyclines, Under-Explored Topics in the Study of Natural Products

## 1.1. INTRODUCTION

Nature has been the source of bioactive compounds used by humans for medicinal purposes since prehistoric times. Whereas these compounds were often used by direct ingestion of the plants, herbs, fungi, or other material they derive from or though crude extracts such as teas or broths, the first isolation of a plant derived natural product, morphine, in pure form was reported in the early 19th century. The field of natural product isolation quickly flourished thereafter, and with it an increased understanding of the ecology of natural products. Deemed "secondary metabolites" at the end of the 19th century, these compounds were differentiated from "primary metabolites," the set of compounds necessary for life, such as amino acids, carbohydrates, and nucleosides. The exact environmental roles which the secondary metabolites play are still not fully appreciated. Yet, these compounds often increase the competitiveness of their producing organisms through a variety of mechanisms. For example, some natural products directly inhibit the growth of other organisms, an activity known as antibiosis.[1-3] The discovery of penicillin, an antibiotic substance produced by a strain of the fungus *Penicillium notatum* garnered much attention as the first antibiotic natural product isolated from a microorganism. However, perhaps the greater legacy of penicillin is its establishment of microorganisms as an important source of bioactive molecules. In particular, the subsequent decades revealed soil-dwelling bacteria of the order Actinomycetales to be prolific producers of natural products. Thousands of antibiotic compounds had been discovered through intensive screening of many bacterial species, since it was thought that the diversity of bacteria drove the diversity of their natural products.[1,4] The advent and

development of molecular cloning techniques led to two key findings which have accelerated and greatly facilitated the discovery of new natural products. First, bacterial natural product biosynthetic genes are now known to be clustered at discrete loci on the genome.[4,5] For the first time, the biosynthetic pathway of natural products could be directly correlated with genes. Cloning of natural product biosynthetic pathways also enabled the possibility of "engineering" the production of their products by relieving transcriptional repression, eliminating shunt pathways leading to less valuable compounds, or increasing the genetic dosage of key biosynthetic enzymes. Secondly, biosynthetic gene clusters are recognized to be modular, where structurally similar natural products are biosynthesized similarly.[5] Thus, since biosynthetic pathways are largely conserved, the ability of a bacterium to produce a given type of natural product could be "screened" genetically, rather than through time-consuming screening of growth and fermentation conditions.[5-7]

Recent advances in genomic sequencing have revealed these bacteria often harbor many natural product gene clusters, suggesting they have an enormous and largely untapped biosynthetic capacity. It is estimated that we have accessed only a fraction of this diversity, as major limitations to the study of bacterial natural products include the inability to culture most bacteria in the laboratory and the transcriptional "silence" of most biosynthetic gene clusters under laboratory conditions.[5] The availability of a wealth of genomic data and our understanding of gene cluster organization combined with current technology have reduced the sometimes decade-long process of identifying gene clusters to a matter of days, increasingly independent of an axenic cultures of the bacteria harboring these gene clusters.[8-10] While powerful, new algorithms can often predict the product of these clusters, these computational methods can be limited since accurate identification of gene clusters or prediction of their products relies on comparison with gene clusters and natural products already discovered. On the other hand, even in a well-defined cluster with

a known product it is still difficult to predict how some structural characteristics are biosynthesized if they have no known precedent.[5,6,11] Given the staggering number of identified but as of yet unclassified bacterial biosynthetic gene clusters, without further work to assign functions to uncharacterized genes, these limitations will continue to impede the full utility of the massive number of bacterial genomes sequenced.[5,12,13]

## 1.2. BIOSYNTHESIS OF THE PEPTIDYL NUCLEOSIDE ANTIBIOTICS

### *1.2.1. Introduction*

The peptidyl nucleoside antibiotics (PNA) comprise a subset of the nucleoside natural products whose study exemplifies the challenges associated with natural products lacking structural conservation with compounds from other mode widely studied groups of natural products.[14] Minimally, the PNA are composed of a nucleoside with at least one pendant amino acid attached via a peptide bond. However, given the diversity of PNA structures, this group of compounds can be best conceptualized as three components: the nucleobase, the "core" saccharide, and the appended amino acid(s). Frequently, each of these components include structures rarely seen, e.g., hypermodified nucleobases, higher-carbon sugars, and non-proteinogenic amino acids.[14-17] Representative PNA are shown in Figure 1.1. While the biological activities of the majority of the PNA discovered to date are unknown, several PNA have been more closely studied, and their activities include: prokaryotic and eukaryotic protein synthesis inhibition (puromycin[18] (**1**) and blasticidin S[19] (**3**)), prokaryotic tRNA synthetase inhibition (albomycin[20] (**18**), bacterial translocase inhibition (mureidomycins and pacidamycins[21,22] (**5**)), and fungal chitin synthase inhibition (nikkomycins (**15**) and

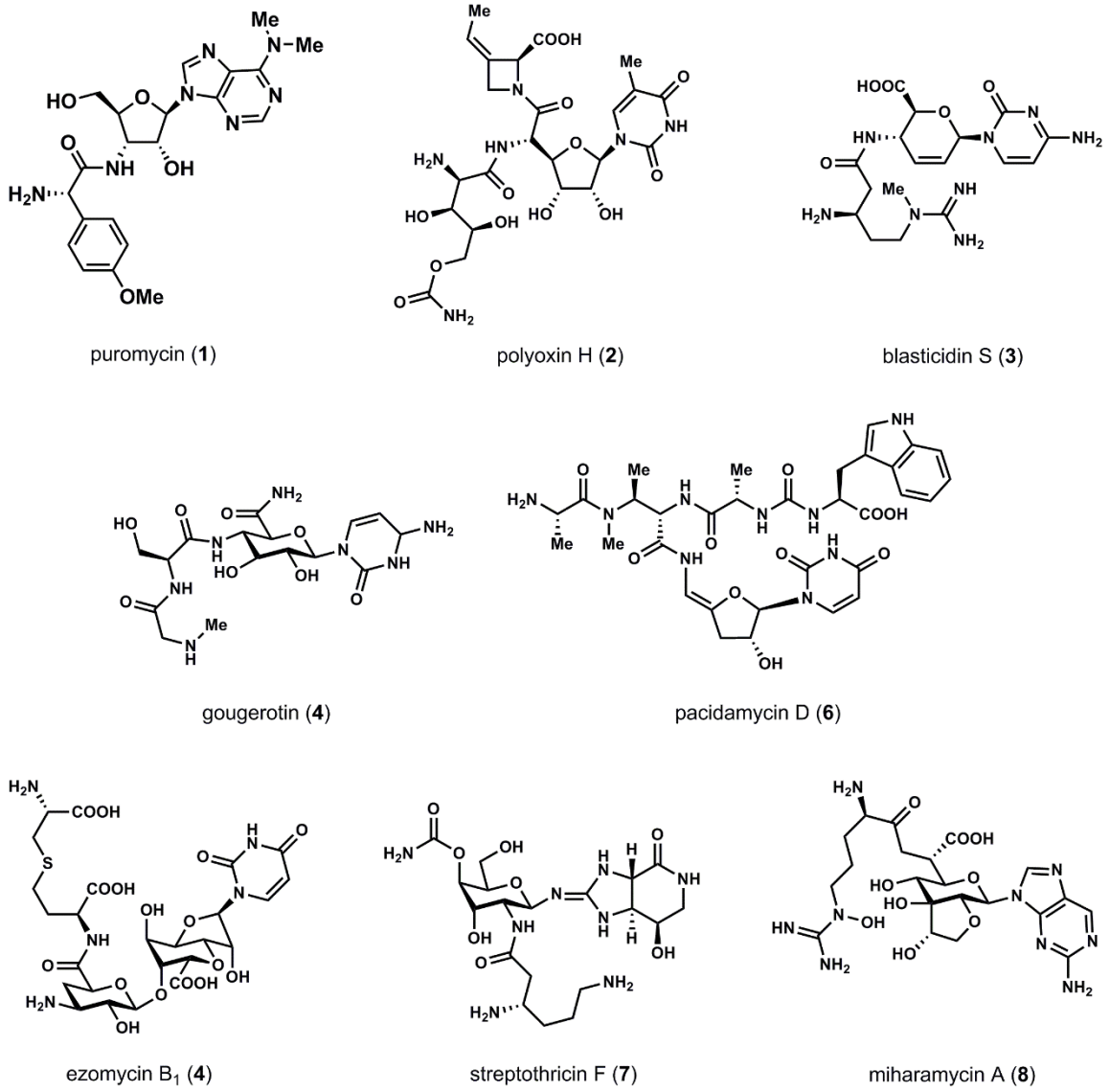



**Figure 1.1:** Structures of some representative peptidyl nucleoside antibiotics

polyoxins[23,24] (**2**)). No PNA are currently used clinically, but nikkomycin Z is under investigation as a first-in-class antifungal agent for the treatment of coccidioidomycosis ("valley fever") in humans.[25-29] Furthermore, puromycin and blasticidin S are heavily used in molecular biology both as selection agents[19] and probes of ribosomal function[18,19]; polyoxin[23] and mildiomycin[30] (**73**) find use as agricultural fungicides; and ningnanmycin

(a stereoisomer of gougeroutin (**4**)) is used as an agricultural viricide.[31] In light of the diverse biological activities exhibited by these "privileged" pharmacophores and unique constituent amino acids and nucleobases, further study of the PNA holds promise for both the development of novel therapeutics and the enrichment of natural product chemistry.

The understanding of the modularity and conservation of gene clusters for many groups of compounds led to rapid discovery of new members of these classes, so a more complete understanding of PNA biosynthesis is expected to greatly aid in the identification of novel PNA gene clusters.[5,12] The structural heterogeneity of the PNA confounds efforts to closely compare individual compounds. However, recently discovered PNA biosynthetic gene clusters have begun to reveal some common biosynthetic motifs, warranting a more useful framework for classification of these compounds. Despite their lack of obviously comparable features, the PNA can be classified based on the nature of their components and the mechanisms employed in their assembly. What follows is a thorough discussion of the biosynthesis of PNA, using only data from currently available biosynthetic gene clusters.

*1.2.2. Amino Acid Attachment*

In spite of having no uniform structural layout, all PNA have at least one pendant amino acid. Investigation into the biosynthesis and attachment of these amino acids can provide a useful "handle" on the biosynthesis of the rest of the molecule.[17,22] A thorough description of the biosyntheses of all unusual amino acids found in the structures of the PNA is beyond the scope of this work, but thus far, only five mechanisms for the establishment of peptide bonds in the PNA have been described. These mechanisms are summarized in Figure 1.2 and can be further classified by whether or not ATP is involved.

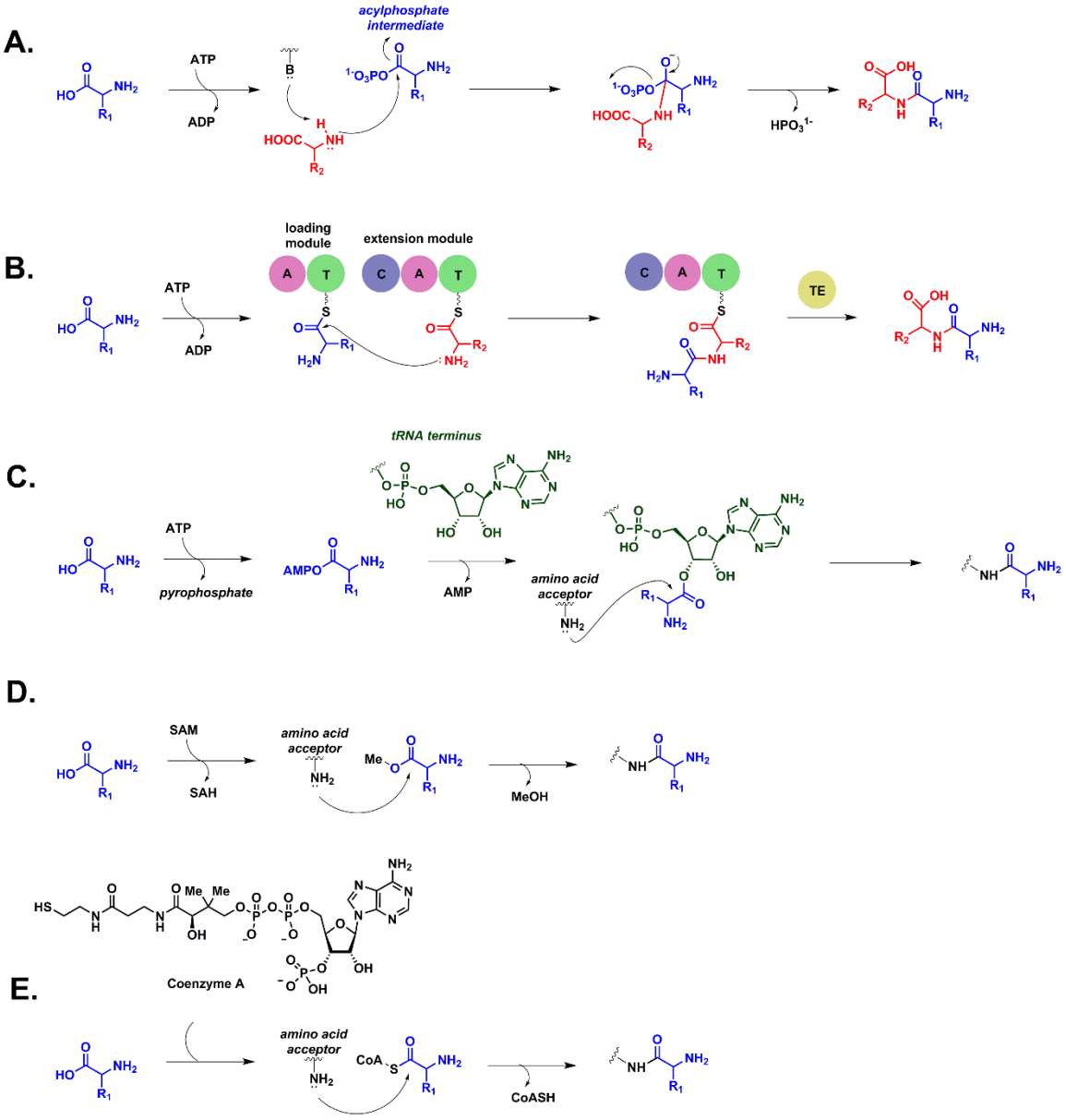



**Figure 1.2:** An overview of peptide bond formation mechanisms utilized in PNA biosynthesis. **A.** Mechanism of ATP-grasp ligase-mediated peptide bond formation; **B.** Mechanism of NRPS-mediated peptide bond formation; **C.** Mechanism of tRNA-mediated peptide bond formation; **D.** Mechanism of methyltransferase-mediated peptide bond formation; **E.** Mechanism of coenzyme A-mediated peptide bond formation

*1.2.2.1. ATP-dependent Peptide Bond Formation*

*1.2.2.1.1. ATP-grasp Ligases*

The ATP-grasp ligases, so-called for their distinct ATP binding motif, comprise a family of enzymes capable of forming an amide bond through activation of a carboxylate via an acyl phosphate intermediate (Figure 1.2A).[32-34] This mechanism is in contrast to activation of carboxylates through adenylation, as observed in adenylate-forming enzymes and non-ribosomal peptide synthetase (NRPS)-associated adenylation domains (Figure 1.2B).[35] Some ATP-grasp ligases make only one amide bond between two substrates, but others are capable of iteratively condensing substrates to create homopolymers.[36-39] While ATP-grasp ligases can be found in primary metabolism, such as glycinamide ribonucleotide synthase (PurD), biotin carboxylase, and D-Ala:D-Ala ligase, many of these ligases have been implicated in various secondary metabolic pathways, either through direct investigation or speculation based on genetic context. Indeed, ATP-grasp ligases are widespread, and their participation in natural product biosynthesis is probably underestimated.[35,37,39]

A hallmark of NRPS-associated adenylation domains is the ability to predict the amino acids they activate *a priori* by extraction of primary sequence residues comprising a "specificity code." This property has been employed in the prediction of the structures of NRPS-based natural products.[40,41] The ATP-grasp ligases lack known indicators of specificity despite selectivity for their particular substrates. The central ATP-binding portion of the ligases is the most well-conserved portion of these enzymes, while the N- and C-terminal sequences display considerable variety.[35] Interestingly, at least one ligase has been demonstrated to be non-specific for the C-terminal residue utilized.[36,42]

*Blasticidin S*

The blasticidin S biosynthetic gene cluster, *bls*, from *Streptomyces griseochromogenes* was found to encode an ATP-grasp ligase, BlsI. Blasticidin S is a hexopyranosyl peptidyl cytosylglucuronic acid (CGA)-derived nucleoside formed through the glycosylation of cytosine with UDP-glucuronic acid with a single amino acid, *N*-methyl-β-arginine, appended to the core saccharide (Figure 1.1).[19] While blasticidin S has only one amide bond, the precise role of BlsI is not known. Methylation of β-arginine is the last step in the biosynthetic pathway of blasticidin S, and demethylblasticidin S is known to be an early intermediate in the biosynthetic pathway, formed through the condensation of β-arginine and cytosinine. However, while BlsI could be responsible for the installation of β-arginine at this step, leucyldemethylblasticidin S is a known later intermediate. Presently, it is not clear whether β-arginine is directly attached to cytosinine by BlsI or if this enzyme is responsible for the attachment of a leucyl-β-arginine dipeptide to cytosinine.[19,43] These possible routes are shown in Figure 1.3. Remarkably, the *bls* cluster includes another putative non-ATP-grasp ligase, BlsK, which will be discussed in section 1.2.3.3. It is possible either a.) one ligase first attaches β-arginine, followed by leucine attachment by the other ligase, e.g., **9 → 10a → 10b** in Figure 1.3; or b.) one of the ligases forms the β-arginine-leucine dipeptide and the other attaches it to cytosinine, e.g., **13 → 10b**.[43] The biosynthesis of arginomycin, a remarkably similar PNA produced by *Streptomyces arginensis* NRRL 15941, likely involves a similar sequence of amino acid attachment, as it, too, is found as a leucyl derivative.[44,45] Accordingly, the arginomycin gene cluster, *arg*, encodes two ligases, ArgJ and ArgK, which are homologs of BlsI and BlsK, respectively.[45] To date, the timing of amino acid attachment in the biosynthesis of arginomycin is not clear.

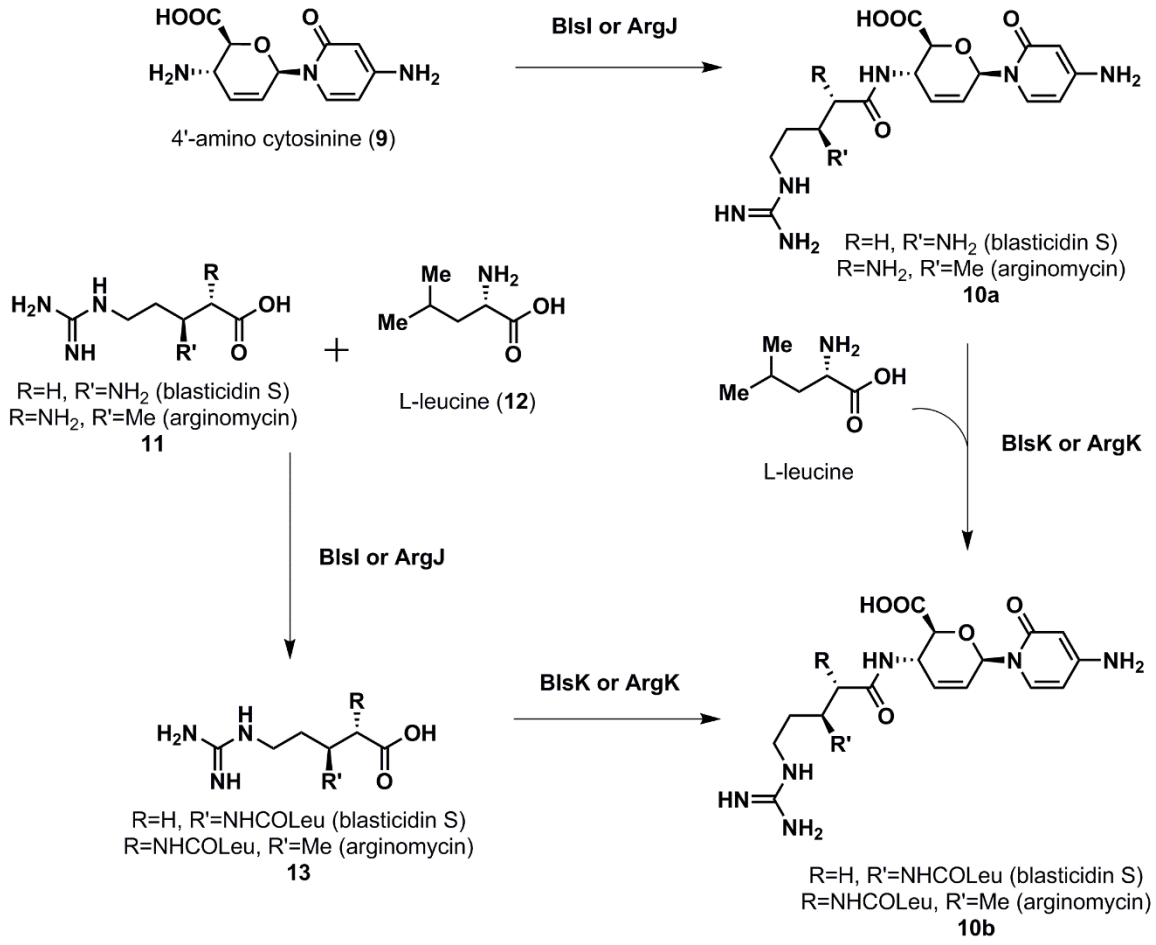



**Figure 1.3:** Possible biosynthetic routes to formation of the leucyl-β-arginine dipeptide found in blasticidin S and arginomycin

*Polyoxins and Nikkomycins*

ATP-grasp ligases are also found in the biosynthetic gene clusters of the polyoxins in *Streptomyces cacaoi* var. *asoensis*, *pol*,[23] and the nikkomycins in *Streptomyces tendae* and *Streptomyces ansochromogenes*, *nik*,[24] and *san*,[46] respectively. The polyoxins and nikkomycins are groups of structurally similar peptidyl nucleosides of an unusual aminohexuronic acid derived from the condensation of phosphoenolpyruvate and a nucleoside monophospate, either UMP, or a derivative thereof (Figure 1.4). The biosynthetic pathway of the polyoxins utilizes 5'-monophosphates of uridine (polyoxin

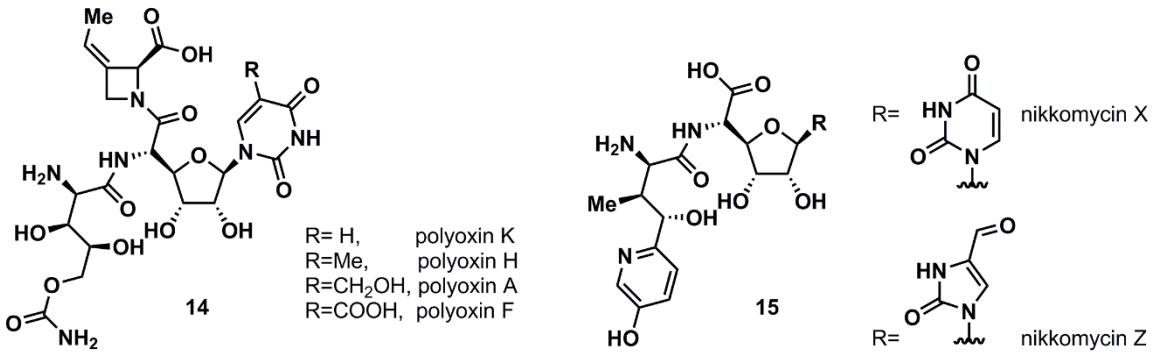



**Figure 1.4:** Comparison of the polyxins (left) and nikkomycins (right) structures

K), 5-methyluridine (polyoxin H), 5-hydroxymethyluridine (polyoxin A), or 5-carboxyuridine (polyoxin F) as a nucleoside substrate.[23] In contrast, the nikkomycins are found as nucleosides of uracil (Z-nikkomycins) or 4-formyl-4-imidazolin-2-one (X-nikkomycins).[24] The biosynthesis of the core aminohexuronic acid saccharide will be discussed elsewhere. The polyoxins have two unusual amino acids attached to the core saccharide, the azetidine-containing polyoximic acid (POIA) and carbamoylpolyoxamic acid (CPOAA), via the amino and carboxylate functionalities, respectively, of the core.[23] Similarly, while the most abundant nikkomycins (the O-type nikkomycins) in culture filtrates are found to have either hydroxypyridylhomothreonine (HPHT) or 2-amino-4-hydroxy-4-(3'-hydroxy-6'-pyridyl) butanoic acid (AHPB) attached to the core via its amino functionality, the minor components nikkomycin I and nikkomycin J, X-type and Z-type nikkomycins, respectively, have an additional glutamate residue appended to the carboxylate of the core saccharide.[24,47-49] While polyoxins and nikkomycins isolated with up to two amino acids appended to their core saccharides, the polyoxins cluster, *pol*, and the two nikkomycins clusters, *nik* and *san*, each encode only a single ATP-grasp ligase, PolG, NikS, and SanS, respectively. The roles of these enzymes in the attachment of the amino acids in these compounds have been substantiated by the interruption of *polG*,[23]

*nikS*,[24] and *sanS*[50] in their respective gene clusters leading to the accumulation of only aminohexuronic acid nucleosides lacking the appended amino acids in the culture filtrates of mutant strains. Thus, unlike other ATP-grasp ligases studied to date, PolG, NikS, and SanS could be remarkably substrate flexible in terms of the their N- and C- terminal substrates.

*Albomycins*

The albomycins gene cluster, *abm*, from *Streptomyces* sp. ATCC 70094, was also found to contain a an ATP-grasp ligase-like protein (NCBI conserved protein domain family PRK000885: "phosphoribosylamine—glycine ligase; provisional"), AbmC. The albomycins are complex PNA that are a conjugate of a ferrichrome siderophore and the unusual 4'-thioxylofuranosyl-based peptidyl nucleoside SB-217452 (**16**).

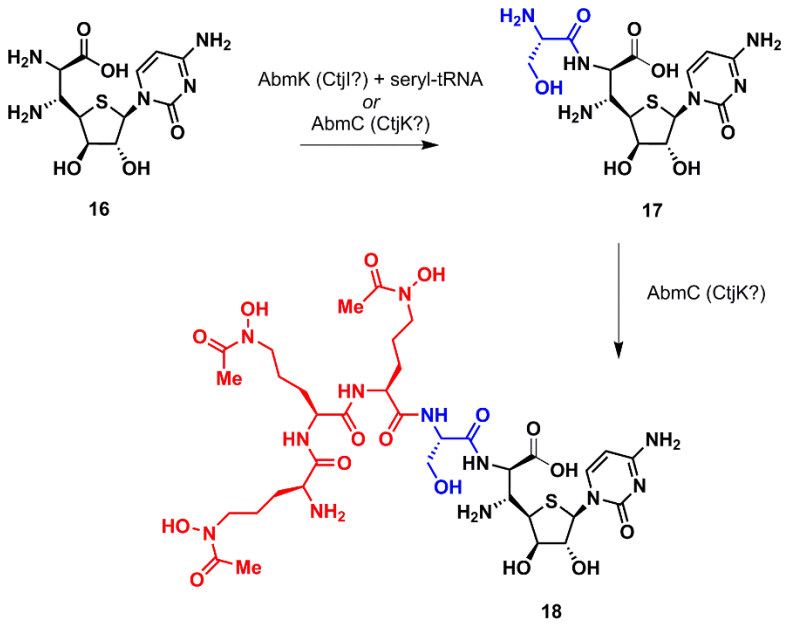



**Figure 1.5:** Hypothetical biosynthetic route to assembly of the three albomycin structural components, the nucleoside (black), serine (blue), and the siderophore (red).

Accordingly, conserved siderophore biosynthetic enzymes are encoded in the *abm* cluster. However, the biosynthesis of the PNA portion is not completely understood.[20,51] It is presently unknown which enzyme(s) are responsible for the attachment of the serine residue found in the structure of the nucleoside portion or for linking the PNA to the siderophore, but given the possibility PolG, NikS, and SanS are able to catalyze the formation of two amide bonds in the biosynthesis of the polyoxins and the nikkomycins, AbmC seems to be a reasonable candidate for both of these reactions (Figure 1.5).

The cloning of the *abm* cluster enabled the identification of a homologous gene cluster, *ctj*, in the publicly-available draft genome sequence of *Streptomyces* sp. *C*. The product of the *ctj* cluster is not presently known, but, as this gene cluster encodes homologues of the enzymes believed to be responsible for the formation of both the 4'-thioxylofuranosyl nucleoside core and siderophore of the albomycins, the encoded product is likely to have a structure similar to the albomycins.[51] Curiously, the *ctj* cluster encodes an overt ATP-grasp ligase, CtjK, instead of an AbmC homolog. Elucidation of the exact roles AbmC and CtjK play will require further work. It should also be noted there is speculation attachment of the serine residue to the nucleoside portion of SB-217452 and likely the Ctj-compound could involve a tRNA-mediated process. In this way, a tRNA synthetase activates serine for another "transferase," or as discussed later for the pacidamycins, the tRNA synthetase serves as the requisite ligase. Indeed, both the *abm* and *ctj* gene clusters encode a seryl-tRNA synthetase, AbmK and CtjI, respectively.[20,51] AbmK was shown to be a mediator of resistance to the action of albomycin on the producer. However, this does not rule out an additional role for AbmK in the attachment of serine.[20]

*A201A*

Recently, an A201A (**21**) biosynthetic gene cluster, *mtd*, was identified in the genome of the sea-dwelling bacterium *Marinactinospora thermotolerans*.[52] A201A shares the $N^6$,$N^6$-dimethyl-3'-amino-3'-deoxyadenosine nucleoside core with the PNA puromycin. However, instead of a tyrosine residue as is found in puromycin, the 3'-amino of the core saccharide is peptide linked to an α-methyl-*p*-coumeric acid disaccharide reminiscent of hygromycin A.[53] The A201A biosynthetic gene cluster, *ata*, was originally found in *Streptomyces capreolus* NRRL 3817, but it was incompletely sequenced, leaving the means of establishing the peptide bond unknown.[53] The *mtd* gene cluster includes the ATP-grasp ligase MtdC. Although it is tempting to assign its role as the necessary peptide synthetase based on the involvement of ATP-grasp ligases in other PNA biosynthetic pathways and the fact MtdC is the only enzyme obviously capable of forming a peptide bond found in the *mtd* gene cluster (Figure 1.6), there are no data to support its role in the biosynthesis of A201A in *M. thermotolerans*.[52]

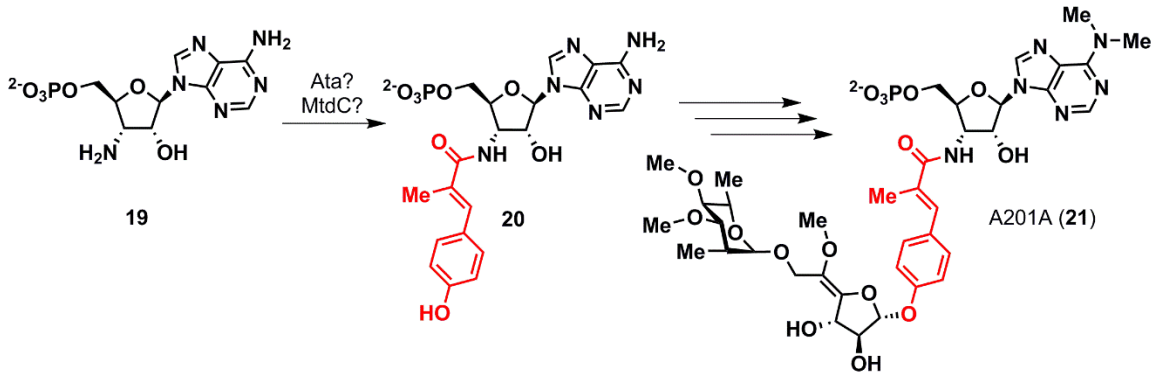



**Figure 1.6:** Hypothetical peptide bond formation in the biosynthesis of A201A

*1.2.2.1.2. Non-Ribosomal Peptide Synthetases*

Non-ribosomal peptide synthetases (NRPS) are renowned for their heavy involvement in the biosynthesis of secondary metabolites. At the heart of NRPS-based peptide synthesis is the use of three basic domains to catalyze three key reactions: a thiolation domain ("T"-domain) which tethers an amino acid substrate via a phosphopantotheinyl prosthetic group, an adenylation domain ("A"-domain) which uses ATP to adenylate an amino acid substrate for loading onto a T-domain, and a condensation domain ("C"-domain) which catalyzes formation of a peptide bond (Figure 1.2B). NRPS enzymes are often found as large, multi-modular enzymes, with each module minimally composed of one of each of the three basic domains, i.e., one T-domain, one A-domain, and one C-domain. Usually, after loading of an initial amino acid "starter unit," each subsequent domain attaches another amino acid, and the lengthened substrate is "passed" to the next domain via a T-domain.[13] As mentioned earlier, A-domain substrates can frequently be predicted on the basis of the domain's primary amino acid sequence, so NRPS-based pathways are considered a form of "templated" biosynthesis.[54,55] In addition to the basic domains, each module may also harbor "tailoring" domains (*cis*-acting domains), such as a methylation domain, cyclization domain, or reductase domain, which act to modify the most recently added amino acid. NRPS systems can also be more complex with non-functional domains, iterative use of modules, and even *trans*-acting tailoring domains.[54-58]

NRPS-based biosynthetic pathways are operative in several PNA, including the structurally similar muraymycins,[59,60] pacidamycins,[21] napsamycins,[61,62] sansamycins,[63] and mureidomycins,[63,64] as well as the streptothricins.[65,66] Of the former group, the NRPS systems of the muraymycins and pacidamycins are the best studied. Given their high structural similarity to the napsamycins, sansamycins, and mureidomycins, the available

data are likely paradigmatic for this class as a whole. With the exception of the streptothricins, these PNA are inhibitors of MraY, the bacterial translocase catalyzing the transfer of phospho-MurNAc-pentapeptide from UDP-MurNAc-pentapeptide to undecaprenyl-phosphate in the biosynthesis of the cell wall component lipid I.[22,67] In contrast, the streptothricins inhibit protein synthesis in both prokaryotes and eukaryotes.[68-71] Remarkably, these disparate biosynthetic pathways all rely on dissociated NRPS systems, where the typically large multi-modular NRPS is replaced with several separate enzymes, with each component comprising no more than three domains.

*Muraymycins*

The muraymycins are a group of PNA isolated from *Streptomces* sp. NRRL 30471. These compounds feature a higher carbon uridyl aminouronic acid nucleoside core common to the caprazamycins and liposidomycins, which are structurally similar *lipo*nucleosides.[59] The biosynthesis of this moiety will be discussed in a later section. Because the muraymycins are also found to be acylated with long alkane side chains, they are often also referred to as nucleoside lipopeptides. The core nucleoside of the muraymycins is further modified through *N*-acylation with a 3-amino-3-carboxypropyl side chain, a modification discussed in section 1.2.3.3, which serves as a "linker" between the core nucleoside and a tripeptide containing an unusual ureido bond. This peptide is matured on the NRPS components and then coupled with the core nucleoside.[59]

Biosynthesis of the peptide chain is expected to begin with the loading of the arginine-derived hexahydro-2-imino-4-pyrimidylglycyl moiety (epicapreomycidine, **22**) onto Mur12, a C-, A-, T-tridomain NRPS module (Figure 1.7). It is then predicted that the C-domain of Mur12 is responsible for establishment of the ureido bond found in the

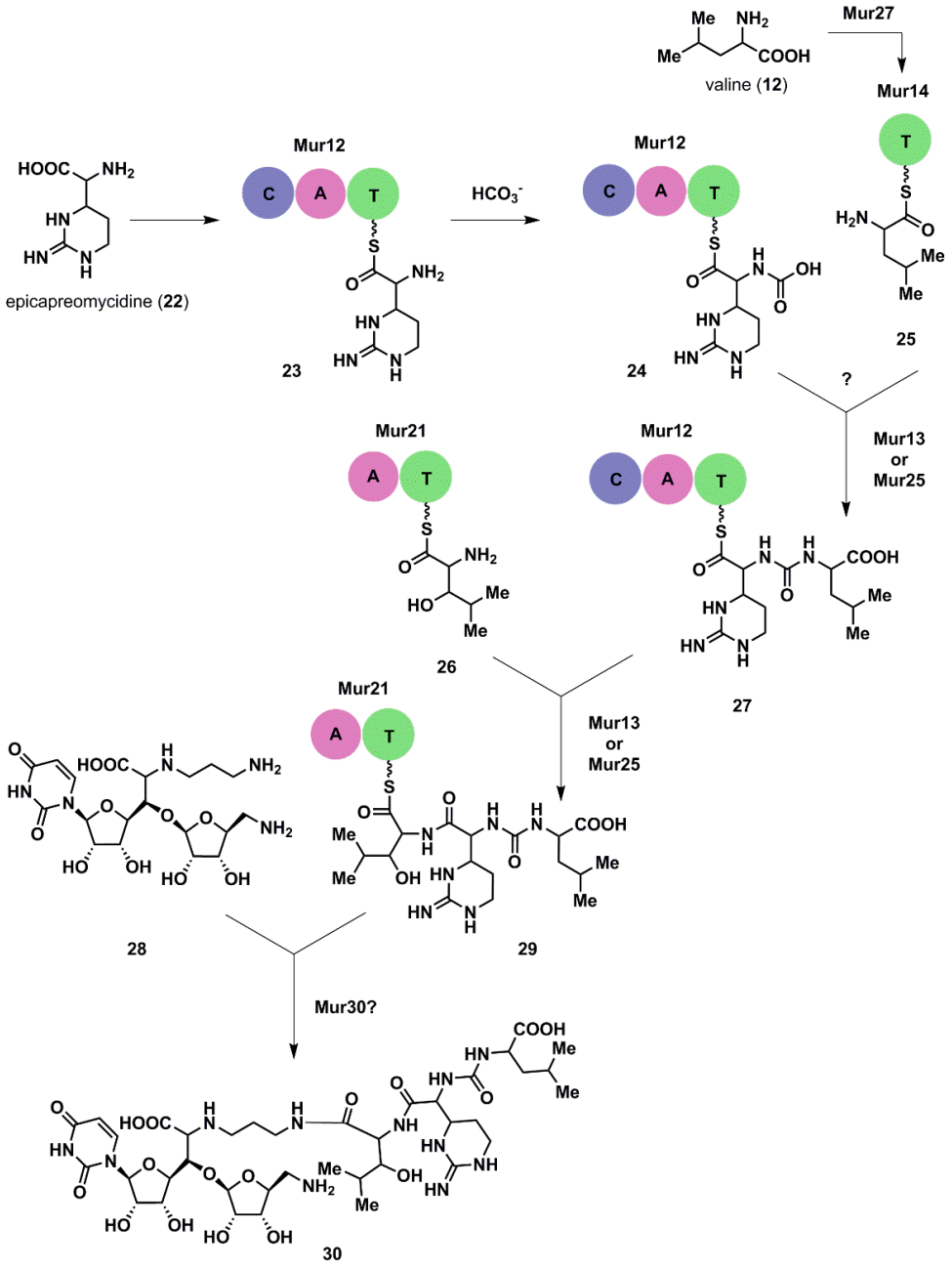



**Figure 1.7:** Proposed NRPS-based biosynthetic pathway of muraymycins

mature peptide by *N*-acylation of epicapreomycidine with bicarbonate. The ureido bond is a unique feature amongst the PNA and is found only in the muraymycins[59] and the pacidamycins.[21,72] Next, Mur14, a standalone T-domain, is predicted to be loaded with

valine by Mur27, and then one of the standalone C-domains, Mur13 or Mur25, could ligate Mur14-tethered valine via its α-amino group to the *N*-carboxy group of Mur12-tethered epicapreomycidine, completing the ureido bond and resulting in a Mur12-tethered pseudo-dipeptide (**27**).[59] Two features make this proposed condensation unique. First, the amide-like ureido bond is formed between an α-amino group and a free carboxylate, unlike the typical NRPS-mediated condensation reactions which involve nucleophilic attack of an α-amine from one amino acid to another amino acid substrate attached to an NRPS T-domain.[73] Secondly, because the new bond does not involve the thioester linkage of a T-domain-tethered amino acid, the Mur12-tethered epicapreomycidine and Mur14-tethered valine would be expected to remain linked via the new ureido bond.[54,55] Yet, the newly linked amino acids are expected to be bound to Mur12. This prediction suggests either a cryptic hydrolytic release mechanism from Mur14 or the involvement of an as-of-yet unidentified hydrolase or thioesterase. Maturation of the peptide chain would be completed through ligation of the Mur12-tethered pseudodipeptide to Mur21-tethered β-hydroxy valine by the remaining standalone C-domain, Mur13 or Mur25. Coupling of the mature peptide (**29**) to the nucleoside core (**28**) is assumed to be accomplished by Mur30, a beta-lactamase-like ligase discussed in another section.

*Pacidamycins*

The pacidamycins are a group of PNA produced by *Streptomyces coeruleorubidus* NRRL 18370 which superficially resemble the muraymycins but lack the acyl chains found in the latter.[21,72] However, like the muraymycins, the pacidamycins feature a uridine-derived core nucleoside, in this case 3'-deoxy-4',5'-enaminouridine, the biosynthesis of which is discussed elsewhere. The core nucleoside is modified through

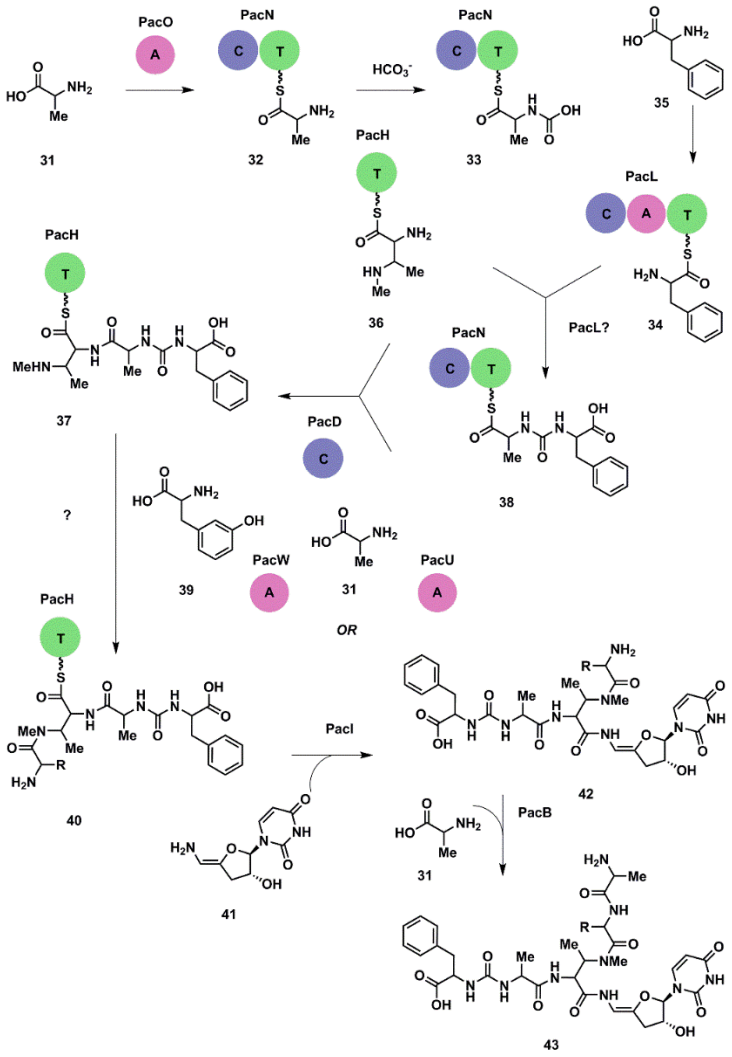



**Figure 1.8:** The NRPS-based pacidamycins biosynthetic pathway

the attachment of a ureido bond-containing tetra- or pentapeptide chain biosynthesized through a dissociated NRPS system. The peptide chain can be visualized as an *Nβ*-methyl 2*S*, 3*S*-diaminoburytic acid (DABA) moiety extended in each direction from its two amino groups by the action of a diverse set of ligases and NRPS components.[74]

Biosynthesis of the peptide chain has been thoroughly investigated through a series of *in vitro* experiments and begins with the auto-loading of PacP, an A-, T-, TE- tridomain

NRPS with desmethyl-DABA (**36**, Figure 1.8).[74-76] Importantly, the TE-domain is expected to be non-functional based on sequence analysis.[74] PacP serves as the scaffold for methylation of desmethyl-DABA by the methyltransferase PacV, and peptide chain extension commences with the transfer of DABA to PacH, a standalone T-domain.[74] Extension from the β-amino group of DABA is facilitated by the action of the standalone A-domains PacU or PacW, which activate alanine or *m*-tyrosine, respectively.[75] Apparently, DABA-tethered PacH and PacU or PacW are sufficient for the attachment of their respective amino acid substrates, suggesting these standalone adenylation domains possess latent peptide synthetase activity. Extension of the DABA α-amino group begins with PacO, a standalone A-domain, loading alanine onto PacN, a C-,T-didomain NRPS module (**32**). Analogously to the proposed role of Mur12 in the biosynthesis of the muraymycins,[59] PacN facilitates the *N*-acylation of the tethered alanine with bicarbonate, ostensibly, but not definitively, through its C-domain (**33**). The ureido bond is made between the new *N*-carboxylate and phenylalanine-tethered PacL, a C-, A-, T-tridomain NRPS module with a predicted non-functional C-domain, and the resultant pseudodipeptide remains linked to PacN (**38**).[74] As with the formation of the ureido bond in the muraymycins, the fine details of the mechanism resulting in the ureido bond in the biosynthesis of the pacidamycins remain unclear. Finally, the PacN-linked ureido dipeptide is condensed with the α-amino group of PacH-linked DABA through the action of the standalone C-domain PacD (**37**), and then the mature peptide is attached to the nucleoside core by the standalone C-domain PacI (**42**).[74] The peptide chain can be further extended after coupling with the nucleoside by the action of PacB (**43**), an unusual ATP-independent ligase which will be discussed later.[75]

*Streptothricins*

The streptothricins have been isolated from at least four organisms: *Streptomyces lavendulae* NBRC 12789*, Streptomyces rochei F20, Streptomyces rochei* NBRC 12908, and *Streptomyces noursei* JA3890b.[65,66,77] These PNA are characterized by a core nucleoside composed of a carbamoylated gulosamine-derived saccharide *C*-glycosidically linked to the purine-like heterocycle streptolidine. The biosynthesis of the core is discussed later.[78] The core nucleoside is decorated with an L-β-lysine oligopeptide chain of up to seven residues, and, while the assembly of this chain and its attachment to the core saccharide rely on a dissociated NRPS system like those discussed earlier, this system is different due to its use of a standalone adenylation domain to perform iterative condensation reactions, as summarized in Figure 1.9.[65]

The streptothricins biosynthetic gene cluster was cloned from *S. rochei* NBRC 12908 and sequenced, revealing the presence of genes encoding NRPS-like components: Orf5 (standalone A-domain), Orf13 (A-,T-didomain NRPS module), Orf18 (T-,C-didomain NRPS module), and Orf19 (standalone A-domain).[65] This finding was in accord with earlier studies in which the streptothricins biosynthetic gene clusters from *S. rochei F20* and *S. noursei* JA3890b were partially sequenced. In particular, Orf5 displayed significant homology to SttA and NpsA, the standalone A-domains previously found in the streptothricins gene clusters in *S. rochei F20* and *S. noursei* JA3890b, respectively.[66] Interestingly, heterologous expression of the *S. rochei* NBRC 12908 gene cluster in which *orf19* was inactivated resulted in the production of only streptothricin F, a streptothricin with only one pendant L-β-lysine residue. The role of Orf19 in the production of longer L-β-lys oligopeptides was further substantiated by the finding that

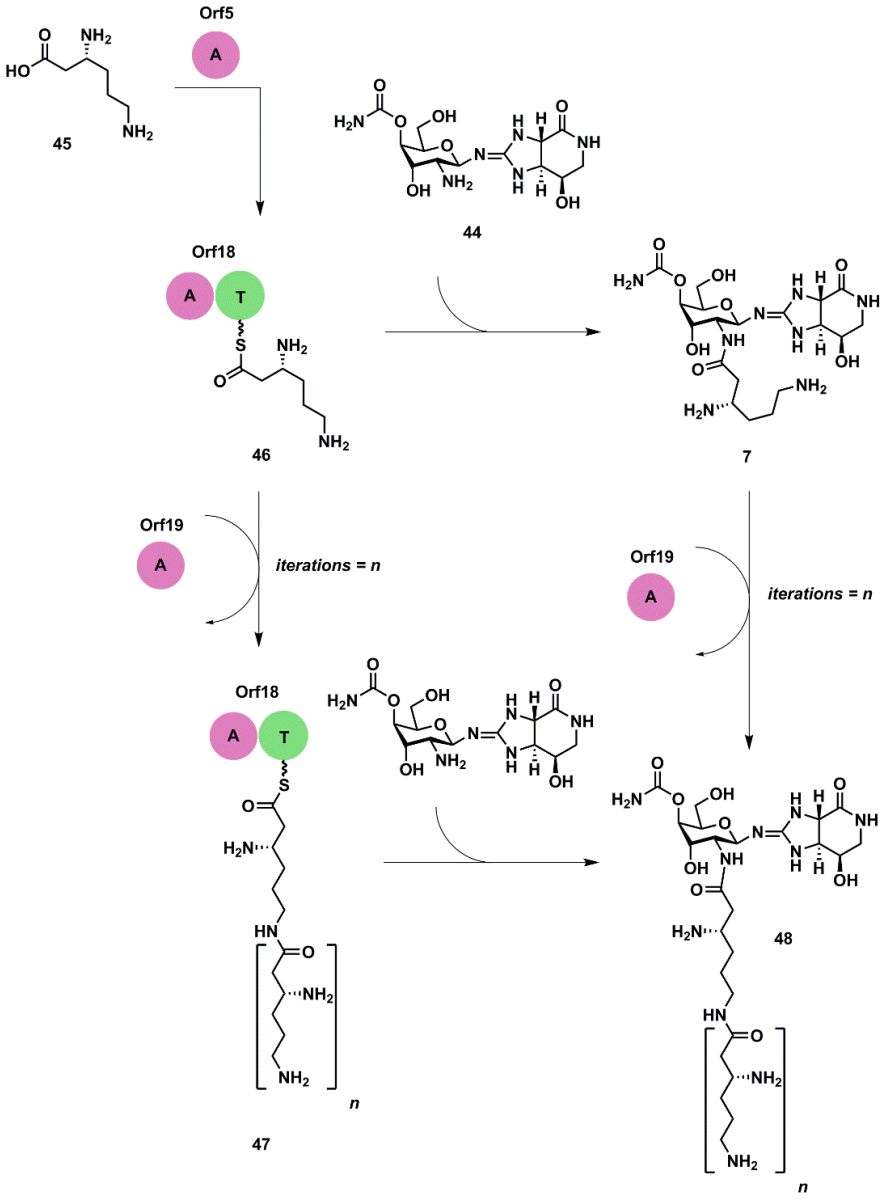



**Figure 1.9:** Streptothricins biosynthetic pathway

the streptothricin gene cluster from *S. lavendulae* NBRC 12789, a strain capable of producing only streptothricin F, lacks a homolog of *orf19*. Furthermore, heterologous co-expression of the *S. lavendulae* NBRC 12789 cluster with *orf19* from the *S. rochei* NBRC

12908 cluster resulted in the production of streptothricins with longer L-β-lysine oligopeptide chains.[65]

A combination of gene deletion experiments and *in vitro* assays led to the construction of a model for the attachment of the L-β-lysine residues to the streptothricin core.[65] Overall, after the loading of the first L-β-lysine residue onto the T-domain of Orf18 by Orf5 (**46**), the oligopeptide can be built on Orf18 through the iterative action of Orf19, i.e., **46 → 48**. Importantly, Orf5 and Orf19 are standalone A-domains with no predictable peptide synthetase activity. The cryptic ability of these A-domains to function as ligases provides support for the role of PacU and PacW, standalone A-domains from pacidamycins biosynthetic pathway,[74,75] to function as ligases as discussed earlier. Orf18 can load a single L-β-lysine onto the streptothricin core to yield streptothricin F (**46 → 7**), or it can load longer polymers onto the streptothricin core (**47 → 48**). Streptothricin F can also serve as a scaffold for the generation of longer L-β-lys polymers through the action of Orf19 (**7 → 48**).

One remaining question is the role of Orf13, an A-,T-didomain NRPS module. Heterologous expression of the *S. rochei* NBRC 12908 gene cluster in which *orf13* was inactivated resulted in the production no streptothricins, indicating a pivotal role for Orf13. However, none of the *in vitro* assays performed seemed to require Orf13, and the pathway leading to production of the L-β-lysine-appended streptothricins was reconstructed without the use of Orf13.[65] This seeming contradiction was never addressed, perhaps indicating the *in vivo* biosynthesis of the streptothricins may differ from the model proposed.

*1.2.2.2. ATP-independent Peptide Bond Formation*

*1.2.2.2.1. N-acylation*

*Amicetins*

Despite the wide use of ATP-dependent ligation mechanisms for the attachment of amino acids in the reported PNA biosyntheses, examples of apparently ATP-independent ligation mechanisms are also known (Figure 1.2E). One such example is the biosynthesis of amicetin-type compounds, as described in Figure 1.10. These PNA are produced by *Streptomyces vinaceusdrappus* and *Streptomyces fasciculatis*, and are composed of a deoxy sugar disaccharide nucleoside of cytosine with a dipeptide of *p*-aminobenzoic acid (PABA) and α-methylserine attached to the cytosine C4-amino group.[79] Amongst the PNA, the site of amino acid attachment in the amicetins is rare, but it is also seen in the septacidin (**49**) and spicamycin (**50**).[80] However, the amicetin group of PNA is the only example whose biosynthetic gene cluster is known.

The amicetin biosynthetic gene cluster was identified in *S. vinaceusdrappus*. Sequence analysis suggested that AmiF, AmiR, and AmiT, which function as a GCN5-related *N*-acetyltransferase, a malonyl CoA-acyl carrier protein transacylase, and an A-, T-didomain NRPS module, respectively, are responsible for assembly of the dipeptide moiety.[79] Deletion of *amiF* in *S. vinaceusdrappus* supported the role of AmiF in the attachment of PABA, as this mutant accumulated the amicetin disaccharide nucleoside lacking amino acids (**53**) and free PABA (**51**). Similarly to the biosynthesis of gougerotin,[31] PABA was expected to require activation with CoA prior to transacylation by AmiF, and AmiL was predicted to be the requisite acyl-CoA synthetase. Accordingly, deletion of *amiL* resulted in accumulation of the same products as the *amiF* deletion

mutant, namely, amicetin nucleoside lacking amino acids and PABA.[79] Unlike gougerotin, which uses CoA to activate both amino acids for attachment,[31] AmiT, an A-,

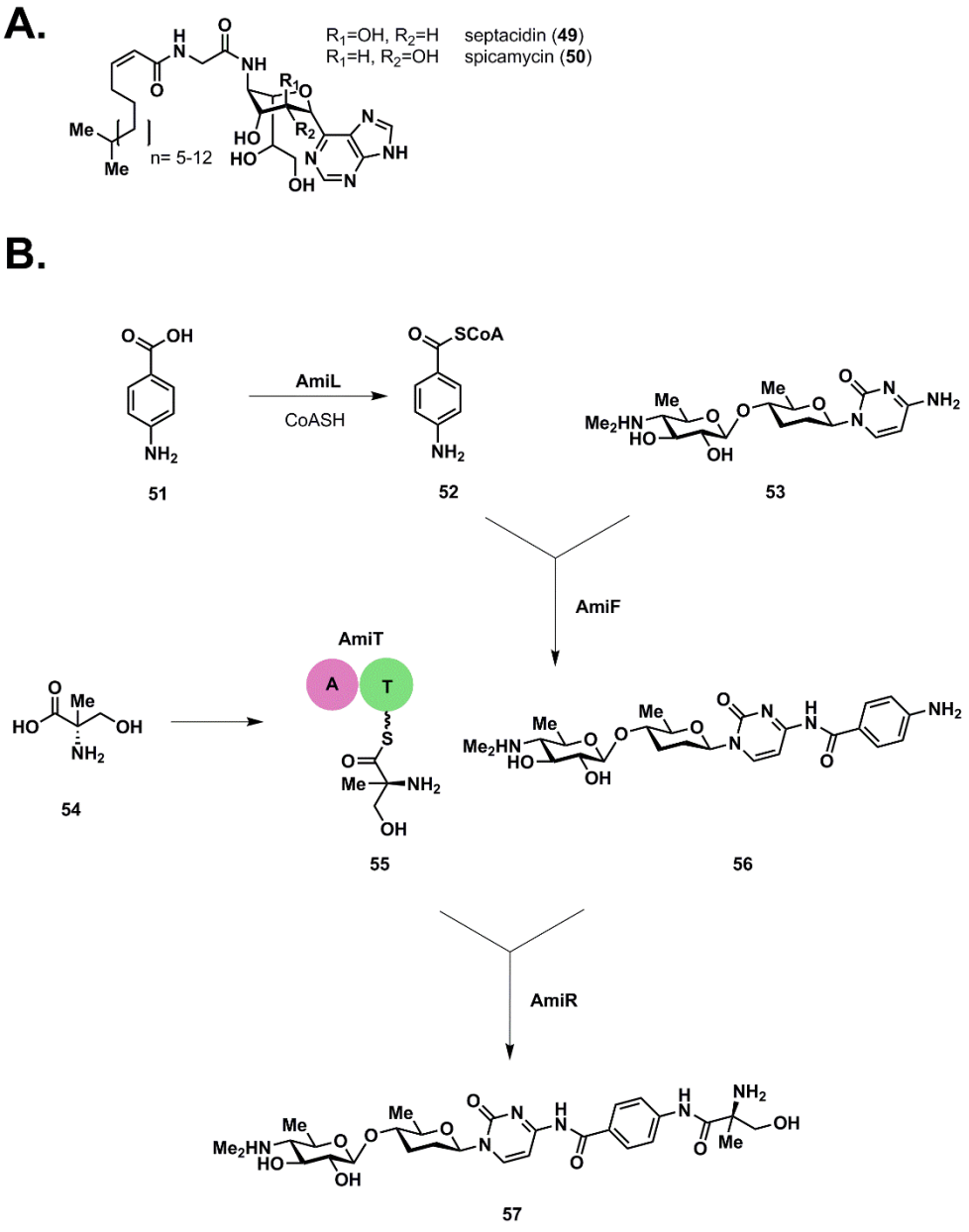



**Figure 1.10: A.** Structures of septacidin and spicamycin; **B.** Amicetins biosynthetic pathway.

T-didomain NRPS module was found to be encoded in the amicetin gene cluster. AmiT was predicted to play a role in the activation of α-methylserine (**54**) for attachment, since the specificity signature of its A-domain indicated substrate preference for serine. Because AmiT features a long N-terminal region with no predicted NRPS modules, the authors suggested this region could harbor some cryptic module, an "X-domain." However, no evidence for this suggestion was provided, and, notably, there seem to be no steps in the biosynthesis of the amicetins requiring unidentified enzymes. The timing for methylation of serine is not known at this time, but it is predicted to occur before loading onto AmiT. Because no other typical ligases, such as those already discussed, appeared to be encoded in the amcetin gene cluster, AmiR, another GCN5-related *N*-acetyltransferase[37] was suspected to serve as the second ligase. This assertion was consistent with the results of the deletion of *amiR*, since amicetins lacking amino acids as well as amicetins bearing only the PABA moiety (**56**) were found to accumulate in this mutant.[79]

*Gougerotin*

*N*-acylation is also used to attach the two amino acids found in gougerotin (**4**), as described in Figure 1.11. This PNA is produced by several organisms,[31] including *Streptomyces graminaerus*, *Streptomyces gougerotii*, *Streptomyces toyocaensis* var. *aspiculamyceticus*, and *Streptomyces* S-514. It is structurally very similar to the bagougeramines.[81] Gougerotin is a hexopyranosyl cytosyl nucleoside with an appended D-serine-sarcosine dipeptide. The gougerotin biosynthetic gene cluster was cloned from *S. graminaerus*, and on the basis of gene deletions and bioinformatic analyses, GouK and GouJ, an acyl-CoA synthetase and an acyl-CoA *N*-acyltransferase, respectively, were proposed to be responsible for installation of the dipeptide chain found in this compound.[31]

GouK is predicted to activate both serine and glycine, the latter of which is *N*-methylated to give sarcosyl-CoA (**66**).

Similarly to blasticidin S biosynthesis,[43] the timing and order of attachment of the two amino acids was not clear. It is not known if D-serine is used, or if L-serine is used and racemized at a later time. For instance, GouJ could attach D- or L-seryl-CoA first

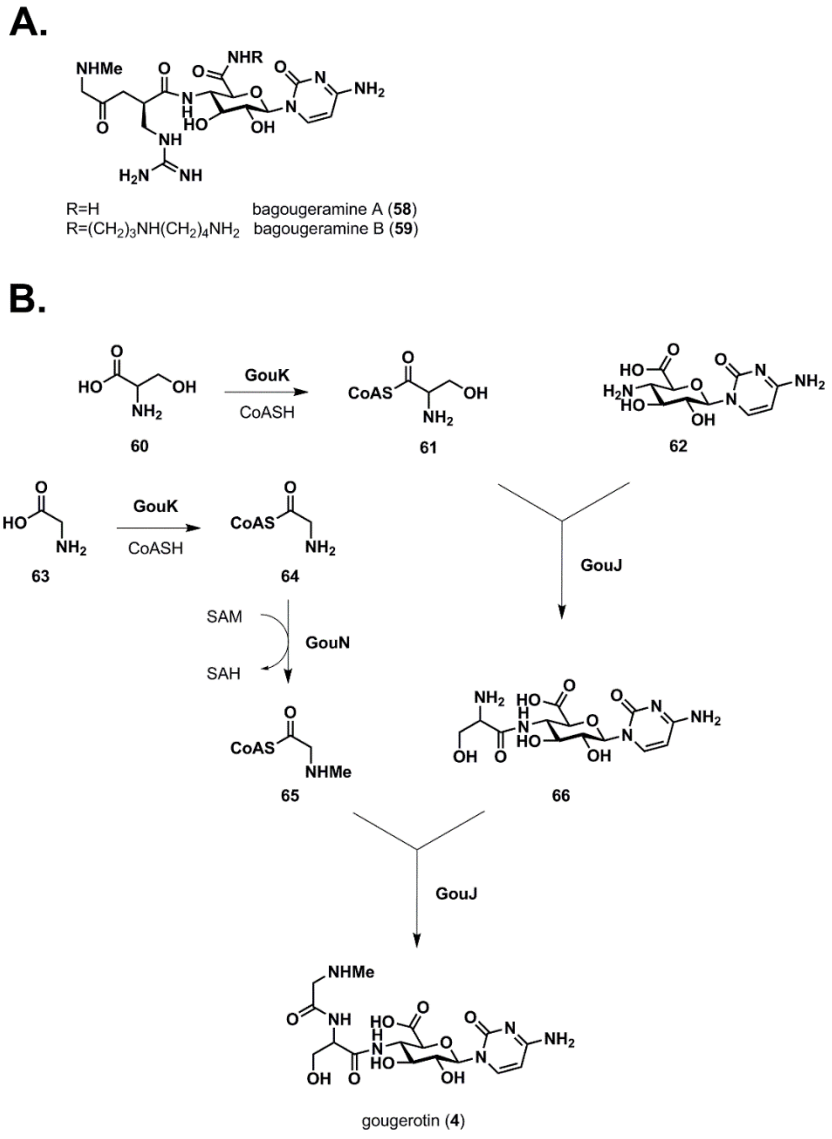



**Figure 1.11: A.** Structures of the bagougeramines; **B.** Gougeroutin biosynthetic pathway.

and then sarcosyl-CoA, or GouJ could mediate the pre-assembly of the two amino acids and then transfer the resultant dipeptide to the nucleoside core structure, with racemization of serine as a "tailoring" step. Evidence for the sequential attachment of serine and then sarcosine was garnered through the heterologous expression of the gougerotin gene cluster with the gene encoding GouN, the glycine *N*-methyltransferase, inactivated. This experiment resulted in the accumulation of only seryl-gougerotin (**66**). Surprisingly, structural analysis of purified gougerotin revealed the presence of ningnanmycin, a previously identified L-serine-containing steroisomer of gougerotin. Thus, it is possible that an unidentified racemase is active in the gougerotin biosynthetic pathway, or GouJ naturally utilizes both isomers of serine.[31] While a similar amino acid ligation mechanism may also be operational in the biosynthesis of the structurally related bagougeramines (**58, 59**), their biosynthetic gene clusters have not yet been identified.

*1.2.2.2.2. Miscellaneous*

*Pacidamycins*

The amino acid ligation methods discussed up to this point have relied on mechanisms to activate the carboxylic acids to facilitate nucleophilic attack by an amine commonly found in metabolic processes. However, the biosynthesis of PNA also provides examples of unusual means to make peptide bonds. First, PacB was recently shown to activate alanine as an aminoacyl-tRNA for attachment to the mature peptide chain of the pacidamycins as a post-coupling tailoring reaction (**42** → **43** in Figure 1.8).[75] PacB showed homology to FemX and FemB from *Weissella viridescens* and *Staphylococcus aureus*, respectively. Since these enzymes utilize tRNA to activate amino acids in peptidoglycan biosynthesis, PacB was predicted to similarly activate alanine using tRNA. Accordingly,

when commercially sourced tRNA and tRNA synthetase were included in reaction mixtures, alanine was efficiently added to the peptide chain of the pacidamycins.

*A-503083*

Another unusual means of peptide bond formation was found in the biosynthesis of the A-503083 group of antibiotics. These natural products are capuramycin-type PNA isolated from *Streptomyces sp.* SANK62799 and *Streptomyces griseus* SANK60196 featuring an amido hexuronyl-uridine nucleoside core. A-503083 A (**70**) is decorated with L-aminocaprolactam (**69**), a cyclized form of L-lysine (Figure 1.12).[82] Analysis of the A-503083 gene cluster from *S. sp.* SANK62799, *cap*, revealed two genes, *capU* and *capV*, encoding a C-, A-, T-containing NRPS and a standalone C-domain, respectively.

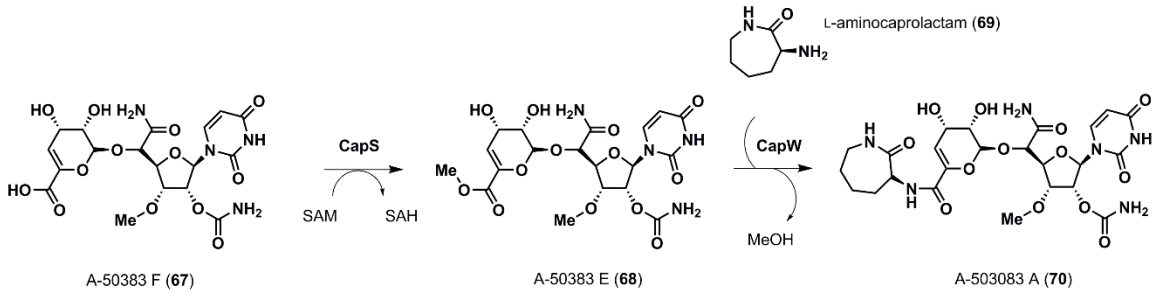



**Figure 1.12:** ATP-independent amide bond formation in the biosynthesis of A-503083 A.

However, the A-domain specificity signature of CapU suggested utilization of L-lysine, which implicated the NRPS enzymes in the biosynthesis of the L-aminocaprolactam moiety. Bordering genes encoding these enzymes is *capW*, an ORF encoding a class C β-lactamase. Because the A-50383A contains a lactam portion, it was believed that CapW could function as an amidase. Since no other candidate exists in the *cap* cluster which could catalyze peptide bond formation, CapW was selected to test whether it can catalyze a peptide bond formation reaction. It is worth mentioning that cultures of *S. sp.*

SANK62799 were found to also produce A-50383 E (**68**), the methyl ester of the core nucleoside, A-503083 F (**67**), which suggested that methylation could also be involved in the activation of the carboxylate group for ligation with L-aminocaprolactam. To examine this possibility, CapW was incubated with the core nucleoside, L-aminocaprolactam, and each of the two methyltransferases found in the *cap* cluster, CapK and CapS. Surprisingly, the reaction containing CapW, CapS, L-aminocaprolactam, and the core nucleoside resulted in attachment of L-aminocaprolactam to the core nucleoside. Formation of the peptide bond was ostensibly accompanied by the elimination of methanol.[35,82] Thus, CapW is the first example of ATP-independent peptide bond formation between a methyl-activated carboxylate and an amine, catalyzed by a β-lactamase homolog. A CapW homolog-encoding gene was also found in the biosynthetic gene cluster of the structurally similar A-500359 group of antibiotics from *Streptomyces griseus* SANK60196.[83]

*Muraymycins*

A β-lactamase may also play a role in the biosynthesis of the muraymycins (Figure 1.7). As discussed in section 1.2.2.1.2., the peptide chain of the muraymycins is matured separately from the nucleoside portion, and Mur30 is predicted to couple the two.[59] Mur30 is a class C β-lactamase of the penicillin binding protein (PBP) transpeptidase superfamily according to analysis using the NCBI Conserved Domain Database. While the Mur30 reaction mechanism is likely similar to that of CapW, the details remain obscure. For example, the CapW catalyzed reaction requires methylation to activate an electrophilic carboxylate,[82] but in muraymycin biosynthesis, the peptide chain is expected to be tethered through a thioester linkage to the T-domain of Mur21 for nucleophilic attack by the nucleoside. To account for this disparity, a cryptic hydrolase may act first to free the peptide chain from Mur21, enabling activation of the free carboxylate by a

methyltransferase. Interestingly, the methyltransferase Mur11 was predicted to play a role in the coupling of the nucleoside and the corresponding peptide chain, but no data have been reported to substantiate this prediction.[59]

*1.2.2.3. Unknown Mechanism*

The gene clusters of blasticidin S (*bls*),[43] mildiomycin (*mil*),[84] and puromycin (*pur*),[18,85] encode putative ligases BlsK, MilH, and Pur6, respectively, which cannot be categorized due to a lack of similarity to any known ligases. Few details regarding their function exist in the literature.

*Blasticidin S and Mildiomycin*

BlsK and MilH were reported to have homology to each other and to lysyl-tRNA ligases, but biochemical evidence for this conclusion based on sequence analysis is lacking, and no further information was available up to date.[45,84] However, mildiomycin biosynthesis in *Streptoverticillum remofaciens* ZJU5119 offers insight into another unique mode of amino acid attachment.[84]

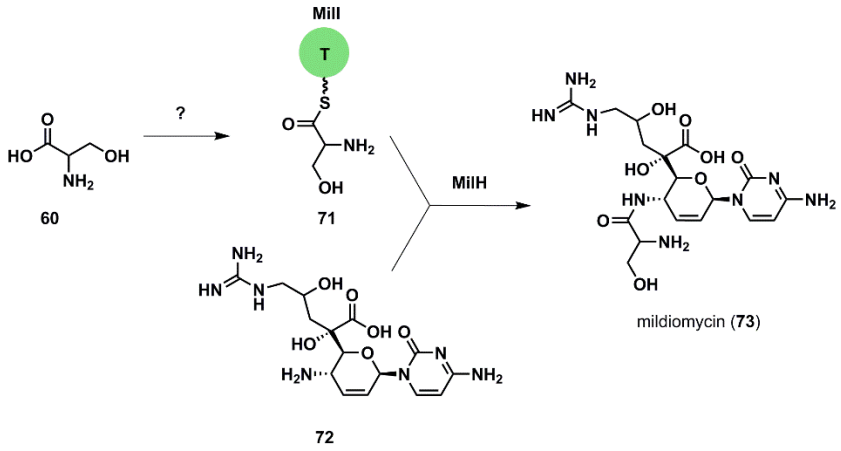



**Figure 1.13:** Peptide bond formation in the biosynthesis of mildiomycin

Mildiomycin is a CGA-derived PNA closely related to blasticidin S,[43] arginomycin,[45] and gougerotin.[31] Similarly to these compounds, it has a 4'-amino group. The amino acid appended at the 4'-amino group is serine. MilH, due to its homology to BlsK, was assigned to perform the ligation reaction. Curiously, the *mil* cluster also includes MilI, a protein with no homology to any proteins in the NCBI database. MilI contains a phosphopantotheinyl (PP) binding site near its N-terminus.[84] The PP prosthetic group is a hallmark of thiolation domains and often acts to tether intermediates in NRPS or PKS pathways.[13,55] Thus, it is predicted that MilI becomes acylated with serine (**71**) prior to the action of MilH.[84] A major point of contention with this model is the *mil* cluster lacks any clear means for the activation of serine for loading onto MilI, such as a standalone adenylation protein or an adenylation domain, as is typical of NRPS enzymology.[55,58]

*Puromycin*

On the other hand, *S. alboniger* Pur6 from the puromycin (**1**) biosynthetic pathway lacks any known ATP-binding motifs, but *in vitro* reactions using cell lysates from recombinant *E. coli* and *Streptomyces lividans* overexpressing *pur6* required ATP and $Mg^{+}$ to ligate L-tyrosine (**73**) to 3'-amino-3'-deoxyadenosine or puromycin aminonucleoside (**72**, $N^6$,$N^6$-dimethyl 3'-amino-3'-deoxyadenosine). The essential role of Pur6 in the biosynthesis of puromycin was also demonstrated by the deletion of *pur6* in *Streptomyces alboniger*, the producing strain, which resulted in ablation of product formation.[86] Clearly more work is warranted to elucidate the mechanisms by which these enzymes catalyze peptide bond formation in their respective biosynthetic pathways.

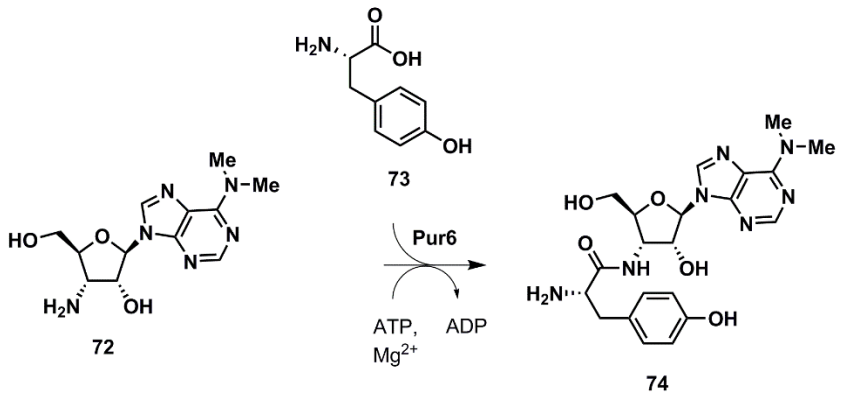



**Figure 1.14:** Pur6-catalyzed peptide bond formation in the biosynthesis of puromycin

*1.2.3. Core Saccharide Origin: Beyond Hexoses Versus Pentoses*

A key outcome of biosyntheses of the core saccharides found in PNA is the installation of functionality amenable to the attachment of pendant amino acids. For the most part, previous surveys of the PNA focused on the number of carbons (e.g., pentose, hexose, or higher) comprising the core saccharide.[15,16] However, this description does not offer significant biosynthetic insight, since analysis of recently elucidated PNA gene clusters indicates varying origins for the core saccharides, even for some with the same ring size.[17,22] Thus, a more detailed classification scheme which considers more than ring size may help to identify similarities between apparently dissimilar PNAs and ultimately aid in the discovery of new PNA gene clusters. We propose the core saccharides can be categorized as *incidental*, *intact*, or *built up*.

*1.2.3.1. Incidental*

The term *incidental* is used to describe core saccharides whose procurement results from incorporation of an intact nucleoside and whose carbon skeletons are not extended. Because the nucleosides themselves are of physiological origin, the pertinent saccharide is ribose. Of the PNA with sequenced biosynthetic gene clusters, puromycin, A201A, and the pacidamycins are the only compounds in this group.

*Puromycin and A201A*

Puromycin and A201A are products of *S. alboniger*[18] and *S. capreolus*,[53] respectively. They share an $N^6$,$N^6$-dimethyl-3'-amino-3'-deoxyadenosine nucleoside derived from adenosine, and the ribose skeleton is not extended any further. Notably, puromycin and A201A are the only purine PNA to date whose biosynthetic gene clusters are known. Even though the *pur* and *ata* clusters share some homologous enzymes, puromycin biosynthesis is the most studied of the two. The biosynthesis of the $N^6$,$N^6$-dimethyl-3'-amino-3'-deoxyadenosine nucleoside begins with oxidation of the 3'-hydroxyl of adenosine triphosphate by Pur10 in *S. alboniger* or AtaP10 in *S. capreolus*.

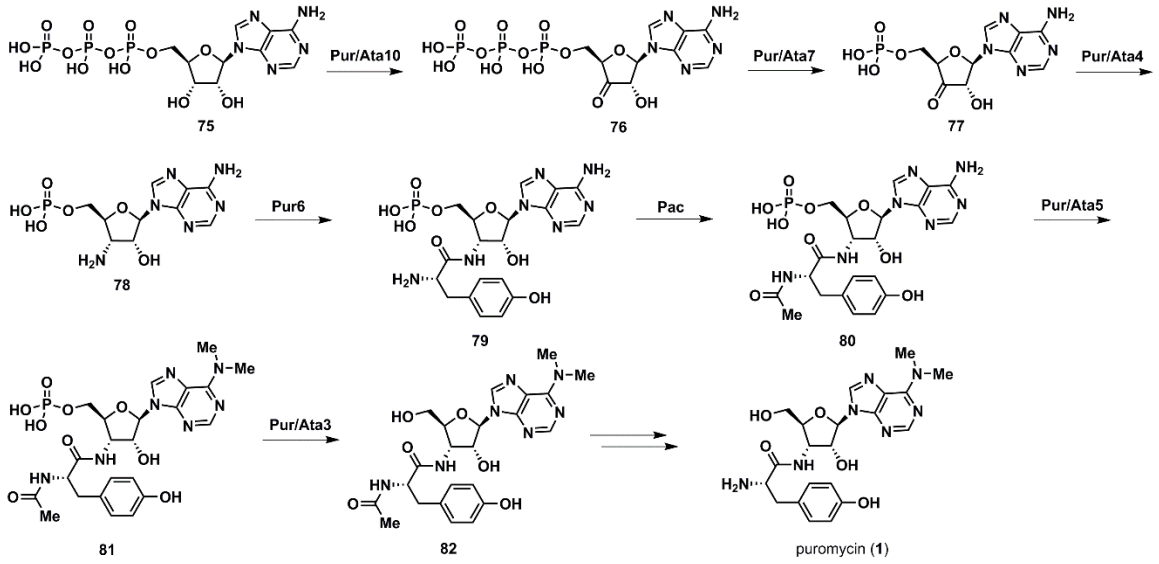



**Figure 1.15:** Biosynthesis of the aminonucleoside core of puromycin; homologous enzymes from the *ata* cluster are noted.

This step is most likely followed by hydrolysis of adenosine triphosphate to adenosine 5'-monophosphate by Pur7 or AtaP7, as 3'-keto adenosine has been shown to be toxic to DNA-dependent RNA polymerase in Gram-negative and Gram-positive bacteria, as well as Erlich acites cells.[18] Then, 3'-transamination is carried out by aminotransferases Pur4

or AtaP4.[18,53] Based on BLAST analyses, these aminotransferases belong to the 3-amino-5-hydroxybenzoic acid synthase family ("AHBA_syn") of fold I, pyridoxal phosphate (PLP)-dependent aspartate aminotransferases. The timing of *N*-dimethylation of 3'-amino-3'-deoxyadenosine has not been definitively established, but it is thought to occur after attachment of tyrosine and modification by an acetyltransferase, Pac. The intermediacy of an acetylated species is supported by an inability of *S. alboniger* cell-free extracts to methylate tyrosyl 3'-amino-3'-deoxyadenosine.[18] Finally, the monophosphatases Pur3 or AtaP3 act to remove the 5'- phosphate. In the biosynthesis of puromycin, the removal of the 5'-phosphate is the penultimate step before export of the compound from the cell,[18] but in the biosynthesis of A201A, the timing of the action of AtaP3 is not clear.[53]

*Pacidamycins*

The pacidamycins from *S. coeruleorubidus* are built upon an unusual 3'-deoxy-ribose-containing uridyl amino nucleoside with an exocyclic alkene.[21,72,87] However, all of the carbon atoms of the core saccharide are contributed by the ribose of uridine, and the core is not further lengthened,[87] characteristics warranting placement of the pacidamycins in the *incidental* group of PNA. As mentioned earlier, although the sansamycins, napsamycins, and mureidomycins are structurally very similar to the pacidamycins, the pacidamycins are the most well-studied of these PNAs, so they are the PNA of focus here.

The first point of contention regarding the origin of the uridine-derived core nucleoside is that uridine is usually found intracellularly as uridine nucleotide.[21,72] Yet, unlike the puromycin[18] or A201A biosynthetic gene clusters,[52,53] the pacidamycins biosynthetic gene cluster contains no phosphatases.[21,72] Given the presence of a 5'-amino group, oxidation of the ribose 5'-hydroxyl is an expected step in the biosynthesis of the

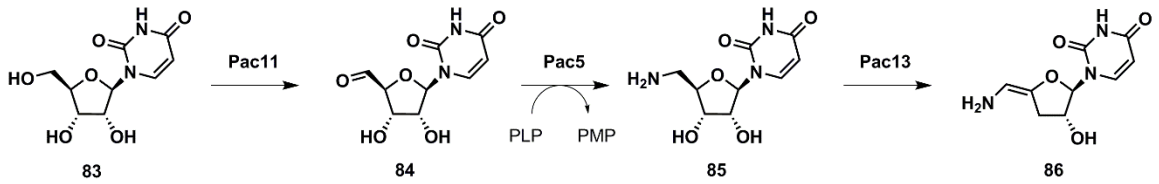



**Figure 1.16:** Biosynthesis of the enamino nucleoside core of the pacidamycins

pacidamycins. Indeed, the *pac* cluster encodes a putative flavoprotein, Pac11, that had been shown to be responsible for oxidizing the 5'-hydroxyl of uridine on the basis of gene deletion and feeding studies in the producing strain, as well as *in vitro* assays (Figure 1.16).[21,87] Oxidation of a ribose 5'-hydroxyl is actually a common step in the biosynthesis of several uridine-derived lipopeptidyl nucleosides, including the caprazamycins, A-90289s, and the capuramycins, but these pathways rely on α-ketoglutarate, non-heme iron-dependent oxygenases that convert uridine 5'-monophosphate to a uridine 5'-aldehyde.[17,22]

The aminotransferase Pac5, a member of the acetylornithine family of the fold I aspartate aminotransferase superfamily, was predicted to perform the 5'-transamination, and this prediction was also borne out by gene deletions and *in vitro* assays.[72,87] No enzymes encoded in the *pac* biosynthetic gene cluster could be assigned to participate in the formation of the unique exocyclic alkene or the 3'-deoxygenation. However, one hypothesis is that spontaneous dehydration eliminates the 3'-hydroxyl, and subsequent tautomerization yields the exocyclic alkene. Surprisingly, the cupin-family protein Pac13 was shown to catalyze both alkene formation and elimination of the 3'-hydroxyl, although the exact mechanism has not been explored. Since Pac13 was unable to generate the enamine nucleoside from 4',5'-saturated pacidamycins, this transformation is expected to occur before coupling with the mature peptide chain.[87]

*1.2.3.2. Intact*

The term *intact* is used to describe those PNA whose core saccharides result from the incorporation of some other non-ribose sugar which is not extended by any additional carbon atoms. These sugars must be activated for condensation with their nucleobases, usually as nucleoside diphosphates. Accordingly, PNA of this group are expected to include a glycosyltransferase gene in their biosynthetic gene clusters.[17] Intact-type PNA with known biosynthetic gene clusters include the cytosylglucuronic acid derivatives blasticidin S, arginomycin, and gougerotin, as well as amicetin and the streptothricins.

*Blasticidin S, Arginomycin, and Gougerotin*

Blasticidin S,[43] arginomycin,[45] and gougerotin[31] are PNA originating from cytosylglucuronic acid (CGA), the condensation product of cytosine and UDP-glucuronic acid. As such, the biosyntheses of these PNAs rely on a glycosyltransferase known as a "CGA-synthase." Mildiomycin is also a CGA-derivative,[84] but its particular biosynthetic pathway precludes designation as an intact-type PNA. Although the abilities of BlsD[88] and MilC,[30,84] the blasticidin S and mildiomycin CGA-synthases, respectively, to glycosylate cytosine with UDP-glucuronic acid have been demonstrated, little work has been done to biochemically characterize CGA-synthases. However, NCBI Conserved Domain Database (CDD) analyses do recognize a "CGA-synthase-like" domain amongst putative CGA-synthases. Another interesting feature of the biosyntheses of these PNAs is the necessity for a nucleoside 2'-deoxyribosyltransferase, because cytosine is rarely found as a free nucleobase intracellularly due to its biosynthesis from uridine.[31,45,84] In support of the need for a nucleoside hydrolase in the biosynthesis of the CGA-derived PNA, supplementation of growing cultures of *S. griseochromogenes* with free cytosine increases the yield of CGA by almost 70-fold.[43] The 2'-deoxyribosyltransferases encoded by the blasticidin S and mildiomycin gene clusters, BlsM[43] and MilB,[30] have been cloned, expressed, and

characterized as *bona fide* cytosine hydrolases. As for typical 2'-deoxyribosyltransferases, these enzymes probably function through a covalent enzyme-deoxyribose intermediate.[43]

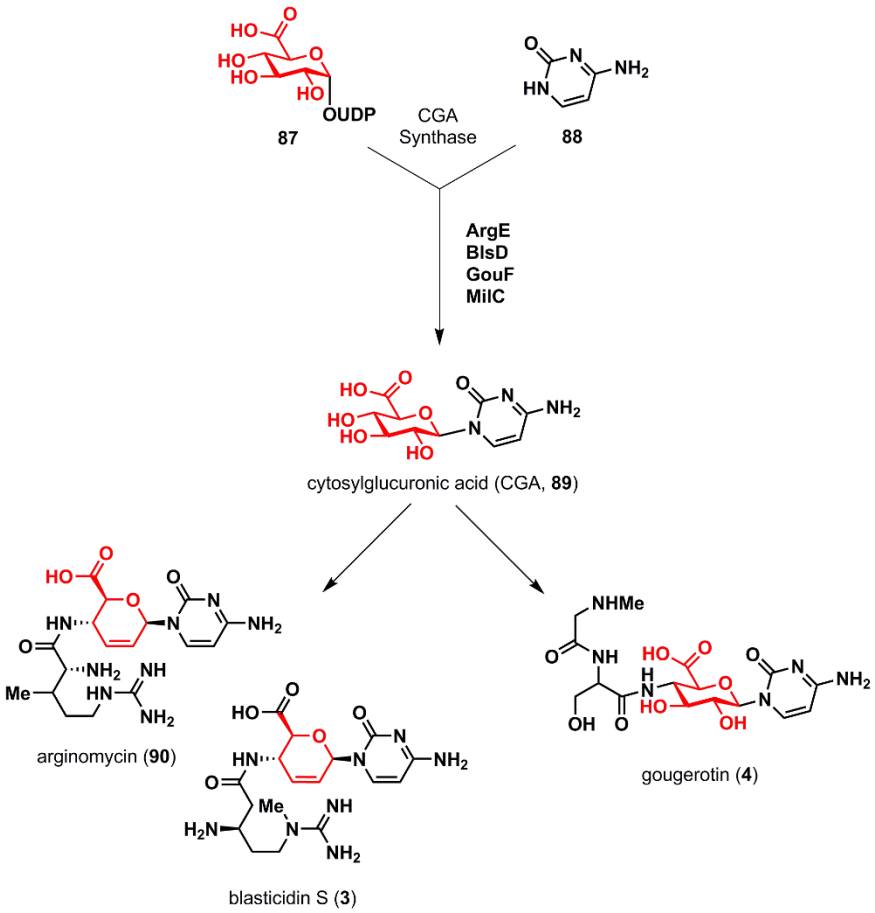



**Figure 1.17:** Biosynthetic origins of the CGA-derivatives blasticidin S, arginomycin, and gougerotin.

Despite sharing the common intermediate CGA with the gougerotin, the PNA blasticidin S,[43] arginomycin,[45] and mildiomycin[84] are actually built upon cytosinine, an unusual unsaturated deoxy form of CGA. The unsaturated glucuronic acid derivative of cytosinine is not well-represented in natural products, but it is reminiscent of an unsaturated hexuronide moiety found in A-503083-type PNA.[82] The biosynthesis of cytosinine has not

been explored, but, in *S. griseochromogenes,* the biosynthesis of cytosinine is related to the production of pentopyranine C and pentopyramine D, which are saturated and are presumably shunt metabolites from the blasticidin S biosynthetic pathway.[89] Comparison of the structures of pentopyranine C, pentopyramine D, and cytosinine led to a hypothetical cytosinine biosynthetic pathway presented by Zhang, et al (Figure 1.18).[89] By analogy to other 3'-deoxy sugars such as CDP-D-ascarylose and GDP-D-colitose, which are cell-well components of several pathogenic Gram-negative bacteria, cytosinine could be derived from 4'-keto CGA.[89,90] This intermediate is also mechanistically relevant considering that the installation of the 4'-amino group found in blasticidin S, arginomycin, mildiomycin, and even gougerotin would require a carbonyl substrate. BlsS, a GMC oxidoreductase homolog, was proposed to oxidize the 4'-position of CGA in the biosynthesis of blasticidin S.[43] While the gougerotin biosynthetic gene cluster in *S. graminareus* includes a D-arabinose dehydratase homolog, GouA,[31] curiously, no homolog of BlsS or GouA is found in the arginomycin[45] or mildiomycin[84] biosynthetic gene clusters.

As expected though, the gene clusters of all four CGA-derived PNA include an aminotransferase of the "AHBA_syn" family of fold I, pyridoxal phosphate (PLP)-dependent aspartate aminotransferase superfamily.[31,43,45,84] It was proposed the blasticidin S aminotransferase, BlsH, a so-called "cytosinine synthase" first generates a Schiff base at the 4'-position of CGA using a pyridoxamine monophosphate (PMP) cofactor instead of PLP,[89] facilitating 4'-3' dehydration via an unsaturated intermediate similar to the $\Delta^{3,4}$-glucoseen intermediate proposed in the mechanism of 3'-deoxygenation of CDP-D-ascarylose.[90] Then, similarly to the 3'-deoxygenation of CDP-

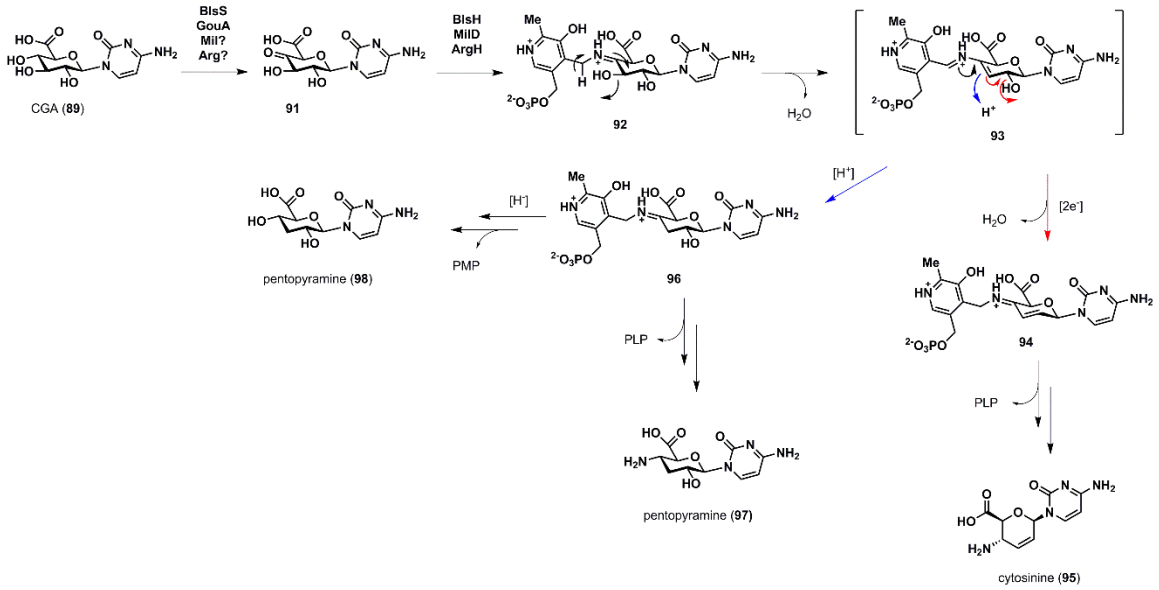



**Figure 1.18:** Proposed PLP-dependent biosynthetic pathway to cytosinine, pentopyranine, and pentopyramine.

D-ascarylose, two-electron reduction of the PMP-sugar adduct and subsequent hydrolysis of PMP could lead to pentopyranine-type products. On the other hand, establishment of a BlsH internal aldimine with the PMP cofactor could result in pentopyramine-type products. Alternatively, reduction of the unsaturated sugar-PMP adduct could lead to elimination of the sugar 2'-hydroxyl, and subsequent establishment of the BlsH internal aldimine with the cofactor would result in cytosinine.[89] Given the high homology between the putative aminotransferases of the CGA-derived PNA, the factors governing the outcome of their reaction with CGA are not clear, especially since gougerotin does not have an unsaturated saccharide.[31]

Another possible biosynthetic route to cytosinine was suggested to involve a radical SAM enzyme by Feng, et al.[91] in the biosynthesis of arginomycin (Figure 1.19).[45] The biosynthetic gene clusters of blasticidin S,[43] mildiomycin,[84] and arginomycin[45] all encode a putative radical SAM enzyme, BlsE, MilG, and ArgF, respectively. BlsE was shown to

be able to decarboxylate CGA *in vitro*,[91] a function now broadly ascribed to homologs of this enzyme, as the NCBI CDD recognizes an "rSAM_BlsE" domain in putative CGA-decarboxylases. Such a function for BlsE is reasonable, as the biosynthesis of pentopyranine C is predicted to involve decarboxylation of CGA.[89,92] However, the conservation of a radical SAM protein in the gene clusters of blasticidin S, arginomycin, and mildiomycin is intriguing, since these PNA all share an unsaturated core saccharide. Furthermore, pentopyranine/pentopyramine-like products have only been isolated from the blasticidin S producing organism.[89,92] The exact role of these radical SAM enzymes in the biosynthetic pathways of these PNAs remains unexplored to date, but may offer insight into the versatility of this class of enzymes.

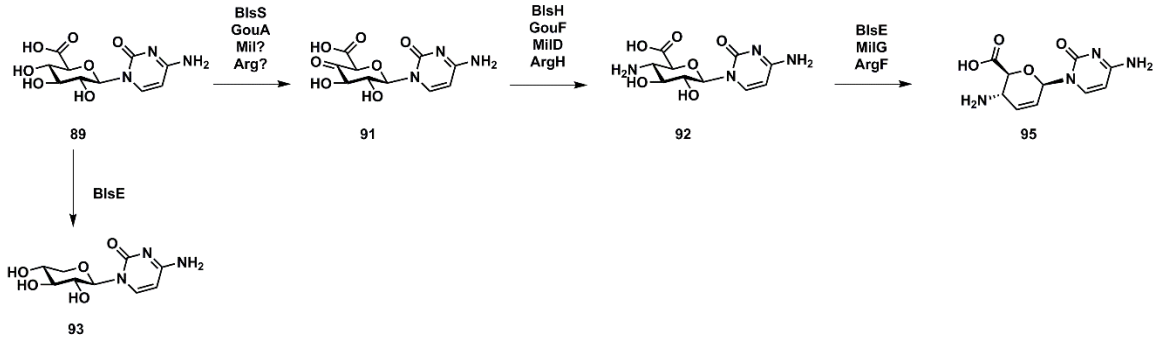



**Figure 1.19:** An alternative, radical-SAM-based biosynthetic route to cytosinine.

*Streptothricins*

The core saccharide of the streptothricins, a carbamoylated gulosamine, was recently found to be derived from UDP-*N*-acetylgalactosamine (UDP-GalNAc).[78] As an intact-type PNA whose biosynthesis involves the incorporation of a non-ribose core saccharide, some means of sugar activation for condensation with its nucleobase is required. Because no NDP-transferase is found in any of the streptothricins gene clusters found to date, and the streptothricin biosynthetic gene cluster is expressed well in

heterologous hosts, a physiologically common sugar was implicated. Previous labeled precursor feeding experiments demonstrated that *N*-acetylglucosamine (GlcNAc) is a precursor. Since UDP-GlcNAc is a component of Gram-positive cell walls, it is, thus, a readily available substrate. The involvement of UDP-GlcNAc in the biosynthesis of the streptothricins nicely reconciled these observations. Unsurprisingly, the requisite glycosyltransferase, StnG (from the streptothricins gene cluster in *Streptomyces* sp. TP-A0356), is distinct from the CGA synthases, and it belongs to glycosyltransferase family GT-2, fold GT-A. While not an absolutely conserved motif amongst GT-A glycosyltransferases,[93,94] StnG has a "DXD" motif which has been associated with the

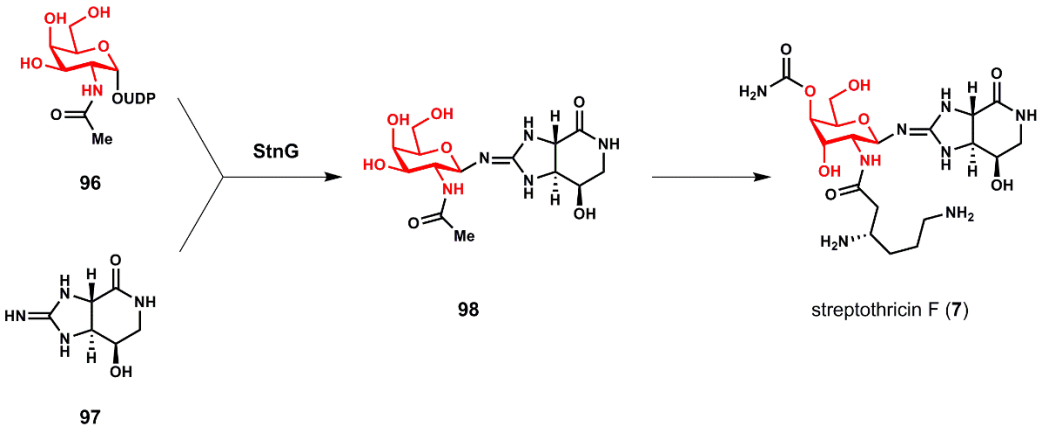



**Figure 1.20:** Origin of the galactosamine core saccharide of the streptothricins.

potential to bind divalent cations.[78] Purified StnG was indeed able to use UDP-GlcNAc to glycosylate the nucleobase streptolidine, but UDP-*N*-acetylgalactosamine (UDP-GalNAc) actually proved to be a better substrate. Since UDP-GalNac is derived from epimerization of UDP-GlcNAc, this result still supported the earlier feeding studies. Epimerization by StnJ and carbamoylation by StnQ complete the biosynthesis of the *N*-acetylgulosamine core saccharide. Notably, the acetyl group appears to act as a "protecting group," since

attachment and elaboration of the poly-β-lysine oligopeptide does not commence until it has been removed by the deacetylase StnI.[78]

*Amicetins*

The biosyntheses of ametcin-type PNA provide another striking example of intact-type PNA, as its core disaccharide is composed of two deoxy sugars biosynthesized "*de novo*" from TDP-glucose.[79] The deoxy sugars amosamine and amicetose are divergently generated from TDP-4'-keto-6'-deoxyglucose following the well-established biosynthetic logic of deoxy sugars,[95,96] and the enzymes necessary for these reactions are all found in the amicetin gene cluster from *S. vinaceusdrappus*. As

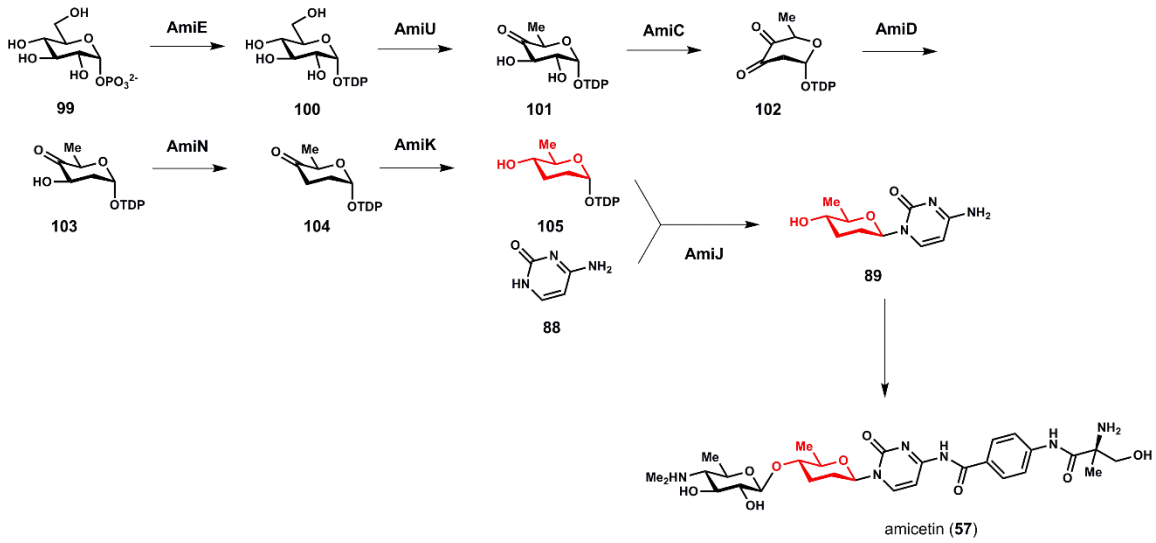



**Figure 1.21:** Deoxy sugar pathway leading to the amicetin core saccharide.

expected, since amicetin-type PNA bear a cytosine base, the amicetin gene cluster also includes a 2'-deoxyribosyltransferase and a CGA synthase, AmiI and AmiJ, respectively. After the coupling of amicetose and cytosine, amosamine is attached by the GT-1, fold GT-B retaining glycosyltransferase AmiG. This glycosyltransferase is apparently substrate

flexible towards the sugar donor, as amicetin derivatives with monomethylated amosamine moieties are also observed. Furthermore, the precise placement of amosamine transfer by AmiG in the amicetin biosynthetic pathway is not known, which may indicate substrate flexibility towards the sugar acceptor.[79]

*1.2.3.3. Built-up*

The last group of PNA are the *built-up* type whose core saccharides may originate from ribose or another sugar through incidental- or intact-type biosynthetic pathways, respectively, but are lengthened or whose cores result from a "piecewise" assembly. The PNA in this group include the polyoxins, nikkomycins, mildiomycin, muraymycin, albomycins, and A-503083s.

*Polyoxins and Nikkomycins*

The polyoxins and nikkomycins are aminohexuronic acids derived from octosyl acid, which, in turn, is known to be derived from phosphoenolpyruvate (PEP) and uridine 5'-monophosphate (UMP).[23,24] The condensation of PEP and UMP in the biosynthesis of the polyoxins and nikkomycins requires the pyruvyl transferases PolA and NikO, respectively. The newly installed 3'-PEP is then transferred to the 5'-position in a reaction sequence that is not well-understood at this point, resulting in the formation of octosyl acid. Importantly, octosyl acid can exist in a "closed chain" perhydrofuropyran form or in an "open chain" form, and the two forms can interconvert, in an ostensibly non-enzymatic process.[15,16] Since the polyoxins and nikkomycins are aminohexuronic acids derived from an octulosuronic acid, a net loss of two carbon atoms must occur at some point. However, virtually nothing is known about how octosyl acid is converted to the hexuronic acid-based core nucleoside of the polyoxins and nikkomycins. The product of this conversion is likely a keto acid, and PolI or NikK are the aminotransferases

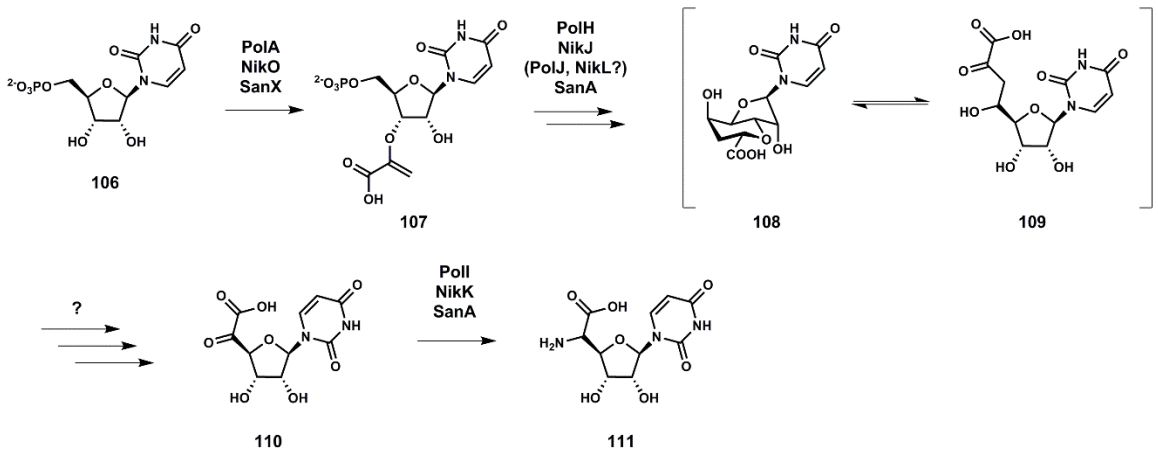



**Figure 1.22:** Biosynthetic route to the built-up core saccharide of the polyoxins and nikkomycins.

predicted to convert the keto hexuronic acid to an aminohexuronic acid. Only the transaminase first half-reaction of NikK has been recently demonstrated.[97] Interestingly, in the *san* cluster, the aminotransferase SanB appears to be fused to a phosphatase at its C-terminus.[46] Octosyl acid isolated from culture filtrates of the polyoxin producer *S. cacaoi* var. *asoensis*, octosyl acid lacks a phosphate group.[15,16] The *nik* and *pol* clusters do encode the monophosphatases NikL[24] and PolJ,[23] respectively, but their placement in the biosynthetic pathways of the nikkomycins and polyoxins has not been established. The composition of SanB may indicate that the phosphatases act after the construction of octosyl acid. Further study of the polyoxin and nikkomycin biosynthetic pathways is expected to provide insight into the unusual chemistry represented therein.

*Mildiomycin*

Mildiomycin is a CGA derivative with a structure similar to those of blasticidin S and arginomycin. Unlike these PNAs though, its cytosinine-based core is effectively lengthened through condensation with α-keto-L-arginine.[84] Feeding studies of the

producing strain *S. remofaciens* demonstrated that all six carbon atoms of $^{13}C_6$- L-arginine are incorporated into mildiomycin. Based on analysis of the mildiomycin gene cluster, MilM and MilN were predicted to attach α-keto-L-arginine to the 4'-amino cytosinine core of this PNA. MilM is an aminotransferase of the "enduracididine-MppP" family of the fold I aspartate aminotransferase superfamily, and this enzyme acts first to deaminate L-arginine. This is a reasonable assignment, as enduracididine is a non-proteinogenic amino acid derived from arginine that is found in several natural products.[98,99] Then, MilN, a dihydropicolinate synthase homolog, could perform the actual condensation reaction. Since all six carbons of α-keto-L-arginine are incorporated into the structure of mildiomycin, the condensation reaction catalyzed by MilN, while not characterized, must proceed through decarboxylation of the 4'-amino cytosinine core nucleoside.[84]

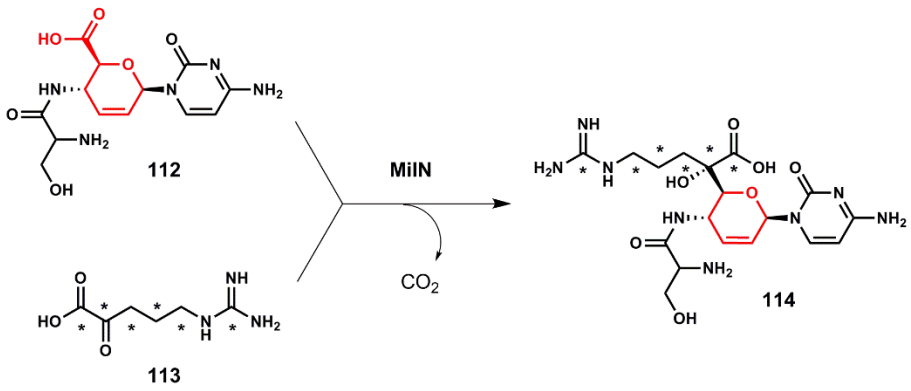



**Figure 1.23:** All six carbon atoms of $^{13}C_6$-L-arginine are incorporated into the mildiomycin core saccharide.

*Albomycins*

The albomycins feature a rare $C_7$-aminothiosugar-based nucleoside, the SB-217452 portion, whose biosynthesis is not understood. The albomycins are a good example of built-up-type PNA, as the aminothiosugar is predicted to be built up from cysteine or

homocysteine and PEP and glycine.[51] Analysis of the albomycins biosynthetic gene cluster led to the proposed roles of the 1-aminocyclopropane-1-carboxylate deaminase homolog AbmD, the radical SAM enzymes AbmM and AbmJ, and the serinehydroxymethyl transferase homolog AbmH in the biosynthesis of the aminothiosugar core saccharide. The resultant 1-phosphate aminothiosugar is then hypothesized to be condensed with cytosine or uracil by AbmG, a deoxyribonucleoside kinase (DK) homolog.

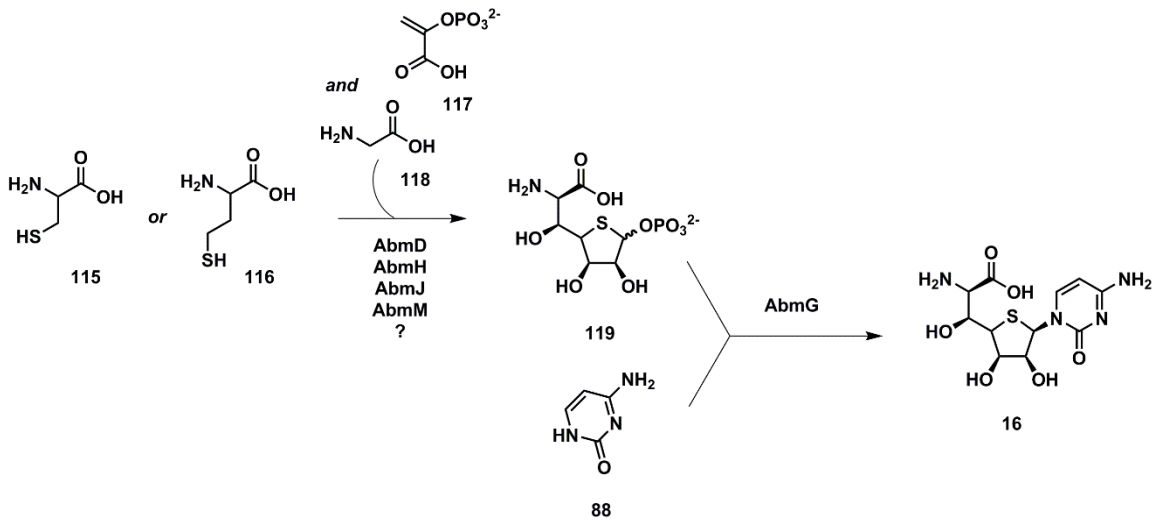



**Figure 1.24:** Overview of albomycin biosynthesis

While DK in general acts to monophosphorylate deoxynucleosides at the 5'-position, human DK has apparently been observed to catalyze phosphorolysis of deoxynucleosides *in vitro* to generate free nucleobases and ribose-1-phosphate, so it was reasoned AbmG could perform the reverse reaction and transfer cytosine or uracil to the aminothiosugar. However, unlike the CGA-based PNA which require the action of a nucleosidase to provide free cytosine,[17,43] the *abm* cluster does not include such an enzyme. Thus, either AbmG can provide both nucleosidase and transfer activities, or a nucleosidase is recruited from elsewhere. Finally, though the timing is not clear, albomycin bearing cytosine is predicted

to be methylated by AbmI to yield albomycin ε and then carbamoylated by AbmE to yield albomycin $\delta_2$. There remain many issues to be resolved in the biosynthesis of the albomycins.

*A-503083s*

Very little is known regarding the biosynthesis of the capuramycinoid A-503083s, but the core saccharide is a six-carbon amide-containing sugar, suggesting a built-up-type biosynthetic pathway.[82,83,100] Feeding studies have suggested that the core is derived from uridine and PEP, and the unsaturated hexuronic acid decorating the core saccharide originates from mannose-1-phosphate. As such, the core saccharide is expected to be composed of eight carbon atoms. Analogous to the polyoxins and nikkomycins, this saccharide is built-up and then partially degraded through an unknown mechanism.

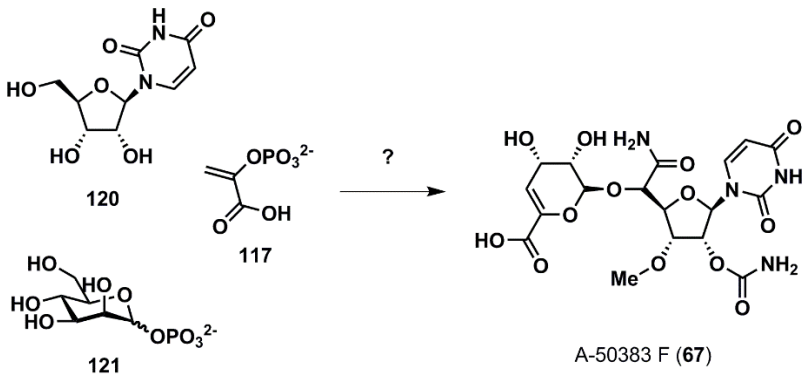



**Figure 1.25:** Putative biosynthetic precursors to A-50383 F

Some insight into the biosynthesis of this core saccharide is afforded by the presence of a serinehydroxymethyltransferase (SHMT) homolog CapH and the non-heme iron(II)-dependent oxygenases CapA and CapD in the A-503083 gene cluster in *S.* sp. SANK 62799. Similarly to the caprazamycins and liposidomycins,[101] the extension of the uridine ribose could involve oxidation of the 5'-hydroxyl to an aldehyde and subsequent transfer

of PEP by the SHMT. The next steps are not clear, and, unlike the gougerotin gene cluster which encodes the amidase GouB,[31] no obvious candidates for formation of the A-503083 amide are found in its gene cluster.

*Muraymycins*

Although the core nucleoside of the muraymycins is derived from an intact molecule of uridine, the ribose is extended through condensation with glycine.[59] Reminiscent of the strategy employed by the caprazamycins and liposidomycins,[101] the 5'-hydroxyl of the core ribose is proposed to be oxidized to an aldehyde by the NAD(P)H-dependent acyl-CoA dehydrogenase homolog Mur22, and then the SHMT homolog Mur17 mediates an aldol condensation with glycine. Unlike the A-503083s, the glycine-modified ribose is not subject to the loss of any carbon atoms. While not necessarily a core extension *per se*, the Mur17 product must modified by the attachment of a 3-amino-3-carboxypropyl linker to accept the amino acid chain of the mature product. This linker is apparently derived from SAM by Mur24. Since Mur24 is predicted to be an aminotransferase by BLAST homology analysis, perhaps the PLP-nucleoside adduct serves as a nucleophile to displace methylthioadenosine from SAM, and hydrolysis of the cofactor would lead to the modified nucleoside. Then, Mur23, a type III PLP-dependent diaminopimelate decarboxylase, decarboxylates the 3-amino-3-carboxypropyl linker. Finally, the core nucleoside is completed by the attachment of 5'-amino-deoxyribose, probably through the action of Mur18, Mur19, and Mur26, a pyrimidine-nucleoside phosphorylase, a GT-1, fold GT-B glycosyltransferase, and a nucleotidyltransferase, respectively.[59]

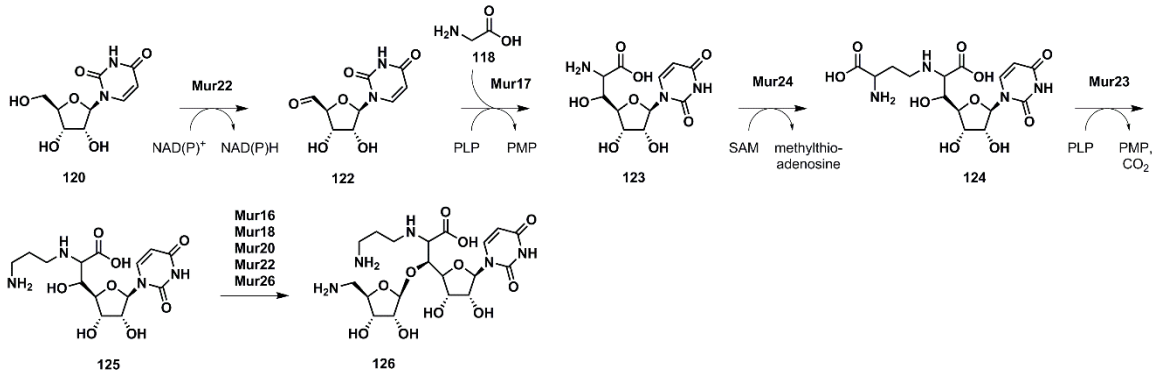



**Figure 1.26:** Proposed biosynthetic pathway of the muraymycins core saccharide.

*1.2.4. Conclusions and Outlook*

Clearly, the PNA offer tremendous opportunity for the study of unique or rarely encountered enzyme mechanisms, as little of the known pathways have been directly characterized. However, deeper study of the PNA is hampered by the novelty of these pathways, since many have almost no enzymological precedent. Currently, the increasing rates of antibiotic resistant pathogens around the world is troubling, and there is mounting pressure to develop novel medicines to stem the tide of infections due to resistant organisms. In light of their diverse biological activities, the PNA are a group of natural products warranting further study. While many of the PNA discovered to date are too toxic for human use, and the growing number of sequenced and annotated genomes will undoubtedly uncover many more PNA gene clusters, the challenges described in this work will likely continue to obfuscate the reliable identification of these clusters. However, by leveraging what is known regarding the biosynthesis of the PNA to offer a more thorough classification of these compounds, it is hoped the field can benefit from the biosynthetic relationships between different pathways by increasing our understanding of PNA gene cluster composition and ultimately improving whole genome sequence-guided efforts to discover novel PNA.

## 1.3. ANTHRACYCLINE AMINO SUGARS

### *1.3.1. Introduction*

Anthracyclines are a class of natural products which are a key component of treatment regimens for several cancers.[102] These compounds are now known to be produced by several bacteria of family Streptomycetaceae, but the first anthracycline was discovered in the culture filtrate from a species of *Streptomyces*. This compound was named rhodomycin probably as a result of its intense red color (*rhodon*= rose colored in Greek). While the spectral characteristics of rhodomycin first suggested the presence of a polycyclic quinone, the structure of rhodomycin was only later found to be a tetracyclic anthraquinone.[103-105] Shortly thereafter, researchers at Farmitalia and Rhone Poulenc Laboratories independently isolated another anthracycline called daunomycin and rubidomycin by each group, respectively. Since daunomycin and rubidomycin were found to be the same compound, it was given a composite name, daunorubicin.[106,107]

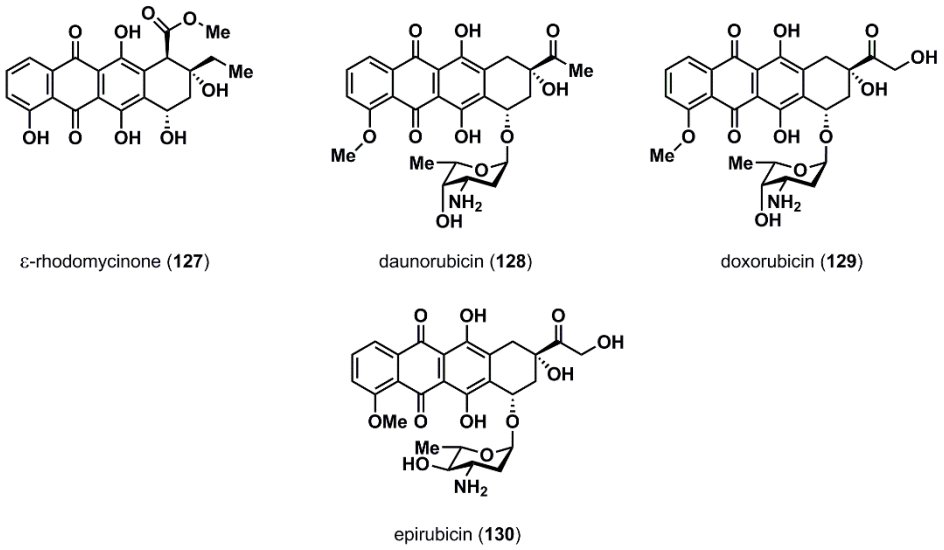



**Figure 1.27:** Structures of representative anthracyclines.

Like rhodomycin, daunorubicin contained a tetracyclic anthraquinone chromophore, but it was also found to be a glycoside, containing a single amino sugar.[108] Very soon after the isolation of the first anthracyclines, they were noted for their poor antimicrobial activity and very potent cytotoxic activity towards several cancer cell lines. The cytotoxicity of these compounds appeared to be related to their propensity to form complexes with nucleic acids, particularly DNA, and the glycoside daunorubicin was found to be much more potent than rhodomycin. Surprisingly, removal of the amino sugar greatly diminished the cytotoxicity of daunorubicin.[109,110] Clinical trials quickly commenced, and although daunorubicin and its derivative doxorubicin showed promising activity in a number of human cancers, they also displayed considerable and often fatal cardiotoxicity.[107,111-113] Given the planar structure of the anthracyclines and their ability to tightly complex DNA, DNA intercalation is believed to be the main mechanism of action through which these compounds exert their biological activity. However, the redox active nature of anthraquinones in general also lent credence to proposals that the anthracyclines generated free radical species *in vivo*. Subsequent work demonstrated a clear correlation between the propensity of the anthracycline aglycons to undergo facile one-electron reduction and the fatal cardiotoxicity associated with anthracyclines.[114-116] The anthracyclines are now understood to exert their biological activity through several mechanisms of action, including DNA intercalation and topoisomerase II inhibition, as well as other unidentified mechanisms.[106,109,117]

The first X-ray crystal structures of daunorubicin and doxorubicin with oligonucleotides very clearly revealed the critical role of their amino sugars. Daunosamine was found to mediate several minor groove interactions, which explained the need for this sugar for relevant biological activity.[110,118] Given the importance of daunosamine to the tight DNA binding of daunorubicin and doxorubicin, this sugar is sensitive to structural

and conformational changes. For example, the α-configuration of daunosamine leads to anthracyclines which bind DNA with much lower affinity than the typical β-configured daunosamine.[119] More importantly, epimerization of the daunosamine 4'-hydroxyl in the anthracycline epirubicin dramatically reduces the cardiotoxicity of doxorubicin. While the prototypical chemotherapeutic anthracycline doxorubicin has a maximum lifetime exposure limit of 550 mg/m$^2$, epirubicin allows an increased maximum lifetime exposure limit of 900 mg/m$^2$.[120] Although over 1000 anthracycline analogs have been explored, to date epirubicin represents the most successful application of glycodiversification as a means to improve the activity profile of anthracyclines.[121]

*1.3.2. Amino Sugar Biosynthesis*

Glycosylated bacterial natural products are notable for the extensive representation of deoxy sugars. These sugars are found to modulate the function of the compounds they are appended to, and often they are key to the biological activities of many natural products. Hundreds of different bacterial deoxy sugars have been found in nature, but, strikingly, these sugars are derived from only a handful of conserved precursors.[95,96]

Beginning with glucose-1-phosphate, the first step in the biosynthesis of bacterial deoxy sugars is the creation of nucleoside diphosphate (NDP)-activated glucose. All five physiological nucleosides are found to be used in the activation of glucose for deoxy sugar biosynthesis, but secondary metabolites frequently rely on a thymidine diphosphate group to activate glucose. Activation as a nucleoside diphosphate serves two purposes. Firstly, this activation step shunts primary metabolic sugars to deoxy sugar biosynthetic pathways. The enzymes of these pathways often rely on the NDP-portion of deoxy sugar intermediates for recognition. Secondly, NDP activation prepares deoxy sugars for their ultimate purpose, the decoration of compounds through glycosylation. The second step is

invariably the 4,6-dehydration of NDP-glucose. The product of this reaction, NDP-4-keto-6-deoxy glucose is shared 2-deoxy sugar biosynthetic pathways, as it can be further modified or attached to a substrate, or it can undergo 2,3-dehydration. The product of the latter reaction is an unstable diketo sugar, so this step is coupled with another reaction like ketoreduction or transamination. In some pathways, NDP-4-keto-6-deoxy glucose undergoes 3,4 ketoisomerization, which provides access to 3-keto-4-hydroxy sugars.[95,96]

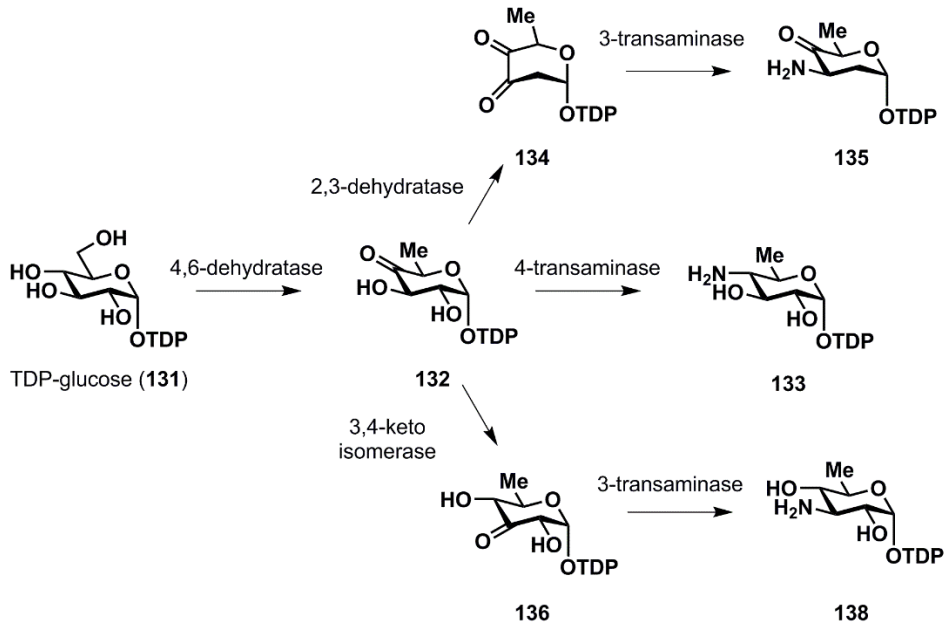



**Figure 1.28:** Deoxy sugar diversity can be generated from only a few conserved intermediates.

Biosynthesis of amino sugars relies on the action of sugar transaminases. These PLP-dependent enzymes catalyze the reversible interconversion of an amino acid and a ketosugar substrate. In the biosynthesis of NDP-amino sugars, sugar transaminases can act at any position of the sugar ring, provided the position is first converted to a carbonyl. Phylogenetic analyses of sugar transaminases may enable a prediction of their sites of

action, but *a priori,* such predictions are not possible given only the amino acid sequence of a particular enzyme. After the installation of an amino group, it can be subject to further modifications as well, including acetylation, methylation, oxidation, carbamoylation, and even glycosylation.[90]

Although daunosamine or epidaunosamine are important determinants of the action of the anthracyclines, very little work has been done to study the biosynthesis of amino sugars in other anthracycline pathways. As a result, many of these amino sugar pathways are predicted but not characterized or validatated.[122] The biosyntheses of amino sugars found in the structures of anthracyclines do not usually feature "exotic" or "unusual" enzyme mechanisms, but a better understanding of these pathways could further glycodiversification efforts in this very useful class of compounds.

*1.3.3. Conclusions and Outlook*

Semi-synthesis is the most commonly employed method to generate anthracycline analogs, but because these molecules are feature many reactive functional groups, synthetic schemes are often lengthy and time consuming. In other cases, anthracyclines are built synthetically, also entailing complicated schemes.[121] Despite the wide utility of the anthracyclines, this class of compounds suffers from dose-limiting, potentially fatal cardiotoxicity[111-113] and the development of resistance in tumors.[117,123] The critical roles played by the deoxy sugars appended to anthracyclines make them a prime target for optimization. Bacterial deoxy sugar biosynthetic pathways are modular and rely on strictly conserved intermediates,[95,96] so efforts to bioengineer the organisms producing these compounds by harnessing their ability to efficiently perform what would be difficult synthetic chemistry steps en route to the formation of these structurally complex molecules

could help to bypass long synthetic schemes and offer a novel approach to the improvement of anthracyclines.

## 1.4. DISSERTATION STATEMENT

This dissertation primarily aims to provide insight into the biosynthesis of complex peptidyl nucleoside antibiotics through investigation of the biosynthesis of the miharamycins and amipurimycin. To a lesser degree, this work also aims to detail preliminary efforts towards the characterization of biosynthetic pathways of the deoxy sugars appended to the unusual anthracycline nogalamycin.

The biosyntheses of the miharamycins, amipurimycin, and nogalamycin were selected as topics of study because they appear to represent substantial deviations from well-studied sugar biosynthetic logic and are, therefore, expected to expand our understanding of nature's biosynthetic repertoire. Chapter 1 furnishes a detailed but concise review of the available recent literature regarding the biosynthesis of peptidyl nucleoside antibiotics and addresses glycosylation of anthracyclines. Chapter 2 describes a bioinformatics approach to the identification of the biosynthetic gene clusters of the complex PNA the miharamycins and amipurimycin and provides analysis of their content. Chapter 3 presents the results of efforts to characterize the biosynthesis of the miharamycins and amipurimycin. Finally, Chapter 4 demonstrates the validation of the proposed biosynthetic pathway of TDP-nogalamine, the amino sugar found in nogalamycin, and offers an approach to the *in vivo* production of C1-hydroxylated anthracyclines.

## 1.5. REFERENCES


(1) Raaijmakers, J. M.; Mazzola, M. Diversity and natural functions of antibiotics produced by beneficial and plant pathogenic bacteria. *Annu Rev Phytopathol.* **2012**, *50*, 403.

(2) Newman, D. J.; Cragg, G. M. Natural products as sources of new drugs over the 30 years from 1981 to 2010. *J Nat Prod.* **2012**, *75*, 311.

(3) Dias, D. A.; Urban, S.; Roessner, U. A historical overview of natural products in drug discovery. *Metabolites.* **2012**, *2*, 303.

(4) Genilloud, O. The re-emerging role of microbial natural products in antibiotic discovery. *Antonie Van Leeuwenhoek.* **2014**, *106*, 173.

(5) Bachmann, B. O.; Van Lanen, S. G.; Baltz, R. H. Microbial genome mining for accelerated natural products discovery: is a renaissance in the making? *J Ind Microbiol Biotechnol.* **2014**, *41*, 175.

(6) Weber, T.; Blin, K.; Duddela, S.; Krug, D.; Kim, H. U.; Bruccoleri, R.; Lee, S. Y.; Fischbach, M. A.; Muller, R.; Wohlleben, W.; Breitling, R.; Takano, E.; Medema, M. H. antiSMASH 3.0-a comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids Res.* **2015**, *43*, W237.

(7) Klementz, D.; Doring, K.; Lucas, X.; Telukunta, K. K.; Erxleben, A.; Deubel, D.; Erber, A.; Santillana, I.; Thomas, O. S.; Bechthold, A.; Gunther, S. StreptomeDB 2.0-an extended resource of natural products produced by streptomycetes. *Nucleic Acids Res.* **2015**.

(8) Lucas, X.; Senger, C.; Erxleben, A.; Gruning, B. A.; Doring, K.; Mosch, J.; Flemming, S.; Gunther, S. StreptomeDB: a resource for natural compounds isolated from Streptomyces species. *Nucleic Acids Res.* **2013**, *41*, D1130.

(9) Henrich, C. J.; Beutler, J. A. Matching the power of high throughput screening to the chemical diversity of natural products. *Nat Prod Rep.* **2013**, *30*, 1284.

(10) Cragg, G. M.; Newman, D. J. Natural products: a continuing source of novel drug leads. *Biochim Biophys Acta.* **2013**, *1830*, 3670.

(11) Overbeek, R.; Olson, R.; Pusch, G. D.; Olsen, G. J.; Davis, J. J.; Disz, T.; Edwards, R. A.; Gerdes, S.; Parrello, B.; Shukla, M.; Vonstein, V.; Wattam, A. R.; Xia, F.; Stevens, R. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res.* **2014**, *42*, D206.

(12) Genilloud, O.; Gonzalez, I.; Salazar, O.; Martin, J.; Tormo, J. R.; Vicente, F. Current approaches to exploit actinomycetes as a source of novel natural products. *J Ind Microbiol Biotechnol.* **2011**, *38*, 375.

(13) Walsh, C. T.; Fischbach, M. A. Natural products version 2.0: connecting genes to molecules. *J Am Chem Soc.* **2010**, *132*, 2469.

(14) Rachakonda, S.; Cartee, L. Challenges in antimicrobial drug discovery and the potential of nucleoside antibiotics. *Curr Med Chem.* **2004**, *11*, 775.

(15) Isono, K. Nucleoside antibiotics: structure, biological activity, and biosynthesis. *J Antibiot (Tokyo).* **1988**, *41*, 1711.

(16) Isono, K. Current progress on nucleoside antibiotics. *Pharmacol Ther.* **1991**, *52*, 269.

(17) Niu, G.; Tan, H. Nucleoside antibiotics: biosynthesis, regulation, and biotechnology. *Trends Microbiol.* **2015**, *23*, 110.

(18) Tercero, J. A.; Espinosa, J. C.; Lacalle, R. A.; Jimenez, A. The biosynthetic pathway of the aminonucleoside antibiotic puromycin, as deduced from the molecular analysis of the pur cluster of Streptomyces alboniger. *J Biol Chem.* **1996**, *271*, 1579.

(19) Cone, M. C.; Petrich, A. K.; Gould, S. J.; Zabriskie, T. M. Cloning and heterologous expression of blasticidin S biosynthetic genes from Streptomyces griseochromogenes. *J Antibiot (Tokyo).* **1998**, *51*, 570.

(20) Zeng, Y.; Roy, H.; Patil, P. B.; Ibba, M.; Chen, S. Characterization of two seryl-tRNA synthetases in albomycin-producing Streptomyces sp. strain ATCC 700974. *Antimicrob Agents Chemother.* **2009**, *53*, 4619.

(21) Rackham, E. J.; Gruschow, S.; Ragab, A. E.; Dickens, S.; Goss, R. J. Pacidamycin biosynthesis: identification and heterologous expression of the first uridyl peptide antibiotic gene cluster. *Chembiochem.* **2010**, *11*, 1700.

(22) Walsh, C. T.; Zhang, W. Chemical logic and enzymatic machinery for biological assembly of peptidyl nucleoside antibiotics. *ACS Chem Biol.* **2011**, *6*, 1000.

(23) Chen, W.; Huang, T.; He, X.; Meng, Q.; You, D.; Bai, L.; Li, J.; Wu, M.; Li, R.; Xie, Z.; Zhou, H.; Zhou, X.; Tan, H.; Deng, Z. Characterization of the polyoxin biosynthetic gene cluster from Streptomyces cacaoi and engineered production of polyoxin H. *J Biol Chem.* **2009**, *284*, 10627.

(24) Lauer, B.; Russwurm, R.; Schwarz, W.; Kalmanczhelyi, A.; Bruntner, C.; Rosemeier, A.; Bormann, C. Molecular characterization of co-transcribed genes from Streptomyces tendae Tu901 involved in the biosynthesis of the peptidyl moiety and assembly of the peptidyl nucleoside antibiotic nikkomycin. *Mol Gen Genet.* **2001**, *264*, 662.

(25) Shubitz, L. F.; Trinh, H. T.; Perrill, R. H.; Thompson, C. M.; Hanan, N. J.; Galgiani, J. N.; Nix, D. E. Modeling nikkomycin Z dosing and pharmacology in

murine pulmonary coccidioidomycosis preparatory to phase 2 clinical trials. *J Infect Dis.* **2014**, *209*, 1949.

(26) Stenland, C. J.; Lis, L. G.; Schendel, F. J.; Hahn, N. J.; Smart, M. A.; Miller, A. L.; von Keitz, M. G.; Gurvich, V. J. A practical and scalable manufacturing process for an anti-fungal agent, Nikkomycin Z. *Org Process Res Dev.* **2013**, *17*, 265.

(27) Shubitz, L. F.; Roy, M. E.; Nix, D. E.; Galgiani, J. N. Efficacy of Nikkomycin Z for respiratory coccidioidomycosis in naturally infected dogs. *Med Mycol.* **2013**, *51*, 747.

(28) Gunasekera, A.; Alvarez, F. J.; Douglas, L. M.; Wang, H. X.; Rosebrock, A. P.; Konopka, J. B. Identification of GIG1, a GlcNAc-induced gene in Candida albicans needed for normal sensitivity to the chitin synthase inhibitor nikkomycin Z. *Eukaryot Cell.* **2010**, *9*, 1476.

(29) Nix, D. E.; Swezey, R. R.; Hector, R.; Galgiani, J. N. Pharmacokinetics of nikkomycin Z after single rising oral doses. *Antimicrob Agents Chemother.* **2009**, *53*, 2517.

(30) Li, L.; Xu, Z.; Xu, X.; Wu, J.; Zhang, Y.; He, X.; Zabriskie, T. M.; Deng, Z. The mildiomycin biosynthesis: initial steps for sequential generation of 5-hydroxymethylcytidine 5'-monophosphate and 5-hydroxymethylcytosine in Streptoverticillium rimofaciens ZJU5119. *Chembiochem.* **2008**, *9*, 1286.

(31) Niu, G.; Li, L.; Wei, J.; Tan, H. Cloning, heterologous expression, and characterization of the gene cluster required for gougerotin biosynthesis. *Chem Biol.* **2013**, *20*, 34.

(32) Fawaz, M. V.; Topper, M. E.; Firestine, S. M. The ATP-grasp enzymes. *Bioorg Chem.* **2011**, *39*, 185.

(33) Kothe, M.; Powers-Lee, S. G. Nucleotide recognition in the ATP-grasp protein carbamoyl phosphate synthetase. *Protein Sci.* **2004**, *13*, 466.

(34) Thoden, J. B.; Kappock, T. J.; Stubbe, J.; Holden, H. M. Three-dimensional structure of N5-carboxyaminoimidazole ribonucleotide synthetase: a member of the ATP grasp protein superfamily. *Biochemistry.* **1999**, *38*, 15480.

(35) Goswami, A.; Van Lanen, S. G. Enzymatic strategies and biocatalysts for amide bond formation: tricks of the trade outside of the ribosome. *Mol Biosyst.* **2015**, *11*, 338.

(36) Kino, K.; Noguchi, A.; Nakazawa, Y.; Yagasaki, M. A novel l-amino acid ligase from bacillus licheniformis. *J Biosci Bioeng.* **2008**, *106*, 313.

(37) Iyer, L. M.; Abhiman, S.; Maxwell Burroughs, A.; Aravind, L. Amidoligases with ATP-grasp, glutamine synthetase-like and acetyltransferase-

like domains: synthesis of novel metabolites and peptide modifications of proteins. *Mol Biosyst.* **2009**, *5*, 1636.

(38) Dawlaty, J.; Zhang, X.; Fischbach, M. A.; Clardy, J. Dapdiamides, tripeptide antibiotics formed by unconventional amide ligases. *J Nat Prod.* **2010**, *73*, 441.

(39) Borisova, S. A.; Circello, B. T.; Zhang, J. K.; van der Donk, W. A.; Metcalf, W. W. Biosynthesis of rhizocticins, antifungal phosphonate oligopeptides produced by Bacillus subtilis ATCC6633. *Chem Biol.* **2010**, *17*, 28.

(40) Stachelhaus, T.; Mootz, H. D.; Marahiel, M. A. The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. *Chem Biol.* **1999**, *6*, 493.

(41) Schneider, A.; Stachelhaus, T.; Marahiel, M. A. Targeted alteration of the substrate specificity of peptide synthetases by rational module swapping. *Mol Gen Genet.* **1998**, *257*, 308.

(42) Kino, K.; Kotanaka, Y.; Arai, T.; Yagasaki, M. A novel L-amino acid ligase from Bacillus subtilis NBRC3134, a microorganism producing peptide-antibiotic rhizocticin. *Biosci Biotechnol Biochem.* **2009**, *73*, 901.

(43) Cone, M. C.; Yin, X.; Grochowski, L. L.; Parker, M. R.; Zabriskie, T. M. The blasticidin S biosynthesis gene cluster from Streptomyces griseochromogenes: sequence analysis, organization, and initial characterization. *Chembiochem.* **2003**, *4*, 821.

(44) Argoudelis, A. D.; Baczynskyj, L.; Kuo, M. T.; Laborde, A. L.; Sebek, O. K.; Truesdell, S. E.; Shilliday, F. B. Arginomycin: production, isolation, characterization and structure. *J Antibiot (Tokyo).* **1987**, *40*, 750.

(45) Feng, J.; Wu, J.; Gao, J.; Xia, Z.; Deng, Z.; He, X. Biosynthesis of the beta-methylarginine residue of peptidyl nucleoside arginomycin in Streptomyces arginensis NRRL 15941. *Appl Environ Microbiol.* **2014**, *80*, 5021.

(46) Liao, G.; Li, J.; Li, L.; Yang, H.; Tian, Y.; Tan, H. Cloning, reassembling and integration of the entire nikkomycin biosynthetic gene cluster into Streptomyces ansochromogenes lead to an improved nikkomycin production. *Microb Cell Fact.* **2010**, *9*, 6.

(47) Lauer, B.; Russwurm, R.; Bormann, C. Molecular characterization of two genes from Streptomyces tendae Tu901 required for the formation of the 4-formyl-4-imidazolin-2-one-containing nucleoside moiety of the peptidyl nucleoside antibiotic nikkomycin. *Eur J Biochem.* **2000**, *267*, 1698.

(48) Bruntner, C.; Lauer, B.; Schwarz, W.; Mohrle, V.; Bormann, C. Molecular characterization of co-transcribed genes from Streptomyces tendae Tu901 involved

in the biosynthesis of the peptidyl moiety of the peptidyl nucleoside antibiotic nikkomycin. *Mol Gen Genet.* **1999**, *262*, 102.

(49) Bruntner, C.; Bormann, C. The Streptomyces tendae Tu901 L-lysine 2-aminotransferase catalyzes the initial reaction in nikkomycin D biosynthesis. *Eur J Biochem.* **1998**, *254*, 347.

(50) Li, Y.; Zeng, H.; Tan, H. Cloning, function, and expression of sanS: a gene essential for nikkomycin biosynthesis of Streptomyces ansochromogenes. *Curr Microbiol.* **2004**, *49*, 128.

(51) Zeng, Y.; Kulkarni, A.; Yang, Z.; Patil, P. B.; Zhou, W.; Chi, X.; Van Lanen, S.; Chen, S. Biosynthesis of albomycin delta(2) provides a template for assembling siderophore and aminoacyl-tRNA synthetase inhibitor conjugates. *ACS Chem Biol.* **2012**, *7*, 1565.

(52) Zhu, Q.; Li, J.; Ma, J.; Luo, M.; Wang, B.; Huang, H.; Tian, X.; Li, W.; Zhang, S.; Zhang, C.; Ju, J. Discovery and engineered overproduction of antimicrobial nucleoside antibiotic A201A from the deep-sea marine actinomycete Marinactinospora thermotolerans SCSIO 00652. *Antimicrob Agents Chemother.* **2012**, *56*, 110.

(53) Saugar, I.; Sanz, E.; Rubio, M. A.; Espinosa, J. C.; Jimenez, A. Identification of a set of genes involved in the biosynthesis of the aminonucleoside moiety of antibiotic A201A from Streptomyces capreolus. *Eur J Biochem.* **2002**, *269*, 5527.

(54) Marahiel, M. A. A structural model for multimodular NRPS assembly lines. *Nat Prod Rep.* **2015**.

(55) Walsh, C. T. Insights into the chemical logic and enzymatic machinery of NRPS assembly lines. *Nat Prod Rep.* **2015**.

(56) Ansari, M. Z.; Yadav, G.; Gokhale, R. S.; Mohanty, D. NRPS-PKS: a knowledge-based resource for analysis of NRPS/PKS megasynthases. *Nucleic Acids Res.* **2004**, *32*, W405.

(57) Lee, T. V.; Johnson, R. D.; Arcus, V. L.; Lott, J. S. Prediction of the substrate for nonribosomal peptide synthetase (NRPS) adenylation domains by virtual screening. *Proteins.* **2015**, *83*, 2052.

(58) Walsh, C. T.; Chen, H.; Keating, T. A.; Hubbard, B. K.; Losey, H. C.; Luo, L.; Marshall, C. G.; Miller, D. A.; Patel, H. M. Tailoring enzymes that modify nonribosomal peptides during and after chain elongation on NRPS assembly lines. *Curr Opin Chem Biol.* **2001**, *5*, 525.

(59) Cheng, L.; Chen, W.; Zhai, L.; Xu, D.; Huang, T.; Lin, S.; Zhou, X.; Deng, Z. Identification of the gene cluster involved in muraymycin biosynthesis from Streptomyces sp. NRRL 30471. *Mol Biosyst.* **2011**, *7*, 920.

(60) McDonald, L. A.; Barbieri, L. R.; Carter, G. T.; Lenoy, E.; Lotvin, J.; Petersen, P. J.; Siegel, M. M.; Singh, G.; Williamson, R. T. Structures of the muraymycins, novel peptidoglycan biosynthesis inhibitors. *J Am Chem Soc.* **2002**, *124*, 10260.

(61) Chatterjee, S.; Nadkarni, S. R.; Vijayakumar, E. K.; Patel, M. V.; Ganguli, B. N.; Fehlhaber, H. W.; Vertesy, L. Napsamycins, new Pseudomonas active antibiotics of the mureidomycin family from Streptomyces sp. HIL Y-82,11372. *J Antibiot (Tokyo).* **1994**, *47*, 595.

(62) Kaysser, L.; Tang, X.; Wemakor, E.; Sedding, K.; Hennig, S.; Siebenberg, S.; Gust, B. Identification of a napsamycin biosynthesis gene cluster by genome mining. *Chembiochem.* **2011**, *12*, 477.

(63) Winn, M.; Goss, R. J.; Kimura, K.; Bugg, T. D. Antimicrobial nucleoside antibiotics targeting cell wall assembly: recent advances in structure-function studies and nucleoside biosynthesis. *Nat Prod Rep.* **2010**, *27*, 279.

(64) Isono, F.; Sakaida, Y.; Takahashi, S.; Kinoshita, T.; Nakamura, T.; Inukai, M. Mureidomycins E and F, minor components of mureidomycins. *J Antibiot (Tokyo).* **1993**, *46*, 1203.

(65) Maruyama, C.; Toyoda, J.; Kato, Y.; Izumikawa, M.; Takagi, M.; Shin-ya, K.; Katano, H.; Utagawa, T.; Hamano, Y. A stand-alone adenylation domain forms amide bonds in streptothricin biosynthesis. *Nat Chem Biol.* **2012**, *8*, 791.

(66) Fernandez-Moreno, M. A.; Vallin, C.; Malpartida, F. Streptothricin biosynthesis is catalyzed by enzymes related to nonribosomal peptide bond formation. *J Bacteriol.* **1997**, *179*, 6929.

(67) Brandish, P. E.; Kimura, K. I.; Inukai, M.; Southgate, R.; Lonsdale, J. T.; Bugg, T. D. Modes of action of tunicamycin, liposidomycin B, and mureidomycin A: inhibition of phospho-N-acetylmuramyl-pentapeptide translocase from Escherichia coli. *Antimicrob Agents Chemother.* **1996**, *40*, 1640.

(68) Takemoto, T.; Inamori, Y.; Kato, Y.; Kubo, M.; Morimoto, K.; Morisaka, K.; Sakai, M.; Sawada, Y.; Taniyama, H. Physiological activity of streptothricin antibiotics. *Chem Pharm Bull (Tokyo).* **1980**, *28*, 2884.

(69) Haupt, I.; Hubener, R.; Thrum, H. Streptothricin F, an inhibitor of protein synthesis with miscoding activity. *J Antibiot (Tokyo).* **1978**, *31*, 1137.

(70) Robinson, H. J.; Graessle, O. E.; Smith, D. G. Studies on the Toxicity and Activity of Streptothricin. *Science.* **1944**, *99*, 540.

(71) Waksman, S. A. Production and Activity of Streptothricin. *J Bacteriol.* **1943**, *46*, 299.

(72) Rackham, E. J.; Gruschow, S.; Goss, R. J. Revealing the first uridyl peptide antibiotic biosynthetic gene cluster and probing pacidamycin biosynthesis. *Bioeng Bugs.* **2011**, *2*, 218.

(73) Stachelhaus, T.; Mootz, H. D.; Bergendahl, V.; Marahiel, M. A. Peptide bond formation in nonribosomal peptide biosynthesis. Catalytic role of the condensation domain. *J Biol Chem.* **1998**, *273*, 22773.

(74) Zhang, W.; Ntai, I.; Bolla, M. L.; Malcolmson, S. J.; Kahne, D.; Kelleher, N. L.; Walsh, C. T. Nine enzymes are required for assembly of the pacidamycin group of peptidyl nucleoside antibiotics. *J Am Chem Soc.* **2011**, *133*, 5240.

(75) Zhang, W.; Ntai, I.; Kelleher, N. L.; Walsh, C. T. tRNA-dependent peptide bond formation by the transferase PacB in biosynthesis of the pacidamycin group of pentapeptidyl nucleoside antibiotics. *Proc Natl Acad Sci U S A.* **2011**, *108*, 12249.

(76) Zhang, W.; Ostash, B.; Walsh, C. T. Identification of the biosynthetic gene cluster for the pacidamycin group of peptidyl nucleoside antibiotics. *Proc Natl Acad Sci U S A.* **2010**, *107*, 16828.

(77) Miyashiro, S.; Ando, T.; Hirayama, K.; Kida, T.; Shibai, H.; Murai, A.; Shiio, T.; Udaka, S. New streptothricin-group antibiotics, AN-201 I and II. Screening, fermentation, isolation, structure and biological activity. *J Antibiot (Tokyo).* **1983**, *36*, 1638.

(78) Guo, Z.; Li, J.; Qin, H.; Wang, M.; Lv, X.; Li, X.; Chen, Y. Biosynthesis of the carbamoylated D-gulosamine moiety of streptothricins: involvement of a guanidino-N-glycosyltransferase and an N-acetyl-D-gulosamine deacetylase. *Angew Chem Int Ed Engl.* **2015**, *54*, 5175.

(79) Zhang, G.; Zhang, H.; Li, S.; Xiao, J.; Zhu, Y.; Niu, S.; Ju, J.; Zhang, C. Characterization of the amicetin biosynthesis gene cluster from Streptomyces vinaceusdrappus NRRL 2363 implicates two alternative strategies for amide bond formation. *Appl Environ Microbiol.* **2012**, *78*, 2393.

(80) Price, N. P.; Furukawa, T.; Cheng, F.; Qi, J.; Chen, W.; Crich, D. Biosynthesis of 4-aminoheptose 2-epimers, core structural components of the septacidins and spicamycins. *J Antibiot (Tokyo).* **2014**, *67*, 405.

(81) Takahashi, A.; Ikeda, D.; Naganawa, H.; Okami, Y.; Umezawa, H. Bagougeramines A and B, new nucleoside antibiotics produced by a strain of Bacillus circulans. II. Physico-chemical properties and structure determination. *J Antibiot (Tokyo).* **1986**, *39*, 1041.

(82) Funabashi, M.; Yang, Z.; Nonaka, K.; Hosobuchi, M.; Fujita, Y.; Shibata, T.; Chi, X.; Van Lanen, S. G. An ATP-independent strategy for amide bond formation in antibiotic biosynthesis. *Nat Chem Biol.* **2010**, *6*, 581.

(83) Funabashi, M.; Nonaka, K.; Yada, C.; Hosobuchi, M.; Masuda, N.; Shibata, T.; Van Lanen, S. G. Identification of the biosynthetic gene cluster of A-500359s in Streptomyces griseus SANK60196. *J Antibiot (Tokyo).* **2009**, *62*, 325.

(84) Wu, J.; Li, L.; Deng, Z.; Zabriskie, T. M.; He, X. Analysis of the mildiomycin biosynthesis gene cluster in Streptoverticillum remofaciens ZJU5119 and characterization of MilC, a hydroxymethyl cytosyl-glucuronic acid synthase. *Chembiochem.* **2012**, *13*, 1613.

(85) Lacalle, R. A.; Tercero, J. A.; Jimenez, A. Cloning of the complete biosynthetic gene cluster for an aminonucleoside antibiotic, puromycin, and its regulated expression in heterologous hosts. *EMBO J.* **1992**, *11*, 785.

(86) Angel Rubio, M.; Barrado, P.; Carlos Espinosa, J.; Jimenez, A.; Fernandez Lobato, M. The pur6 gene of the puromycin biosynthetic gene cluster from Streptomyces alboniger encodes a tyrosinyl-aminonucleoside synthetase. *FEBS Lett.* **2004**, *577*, 371.

(87) Ragab, A. E.; Gruschow, S.; Tromans, D. R.; Goss, R. J. Biogenesis of the unique 4',5'-dehydronucleoside of the uridyl peptide antibiotic pacidamycin. *J Am Chem Soc.* **2011**, *133*, 15288.

(88) Gould, S. J.; Guo, J. Cytosylglucuronic acid synthase (cytosine: UDP-glucuronosyltransferase) from Streptomyces griseochromogenes, the first prokaryotic UDP-glucuronosyltransferase. *J Bacteriol.* **1994**, *176*, 1282.

(89) Zhang, Q.; Gould, S. J.; Zabriskie, T. M. A new cytosine glycoside from Streptomyces griseochromogenes produced by the use in vivo of enzyme inhibitors. *J Nat Prod.* **1998**, *61*, 648.

(90) Romo, A. J.; Liu, H. W. Mechanisms and structures of vitamin B(6)-dependent enzymes involved in deoxy sugar biosynthesis. *Biochim Biophys Acta.* **2011**, *1814*, 1534.

(91) Feng, J.; Wu, J.; Dai, N.; Lin, S.; Xu, H. H.; Deng, Z.; He, X. Discovery and characterization of BlsE, a radical S-adenosyl-L-methionine decarboxylase involved in the blasticidin S biosynthetic pathway. *PLoS One.* **2013**, *8*, e68545.

(92) Gould, S. J.; Guo, J.; Geitmann, A.; DeJesus, K. Nucleoside intermediates in blasticidin S biosynthesis identified by the in vivo use of enzyme inhibitors. *Canadian Journal of Chemistry.* **1994**, *72*, 6.

(93) Coutinho, P. M.; Deleury, E.; Davies, G. J.; Henrissat, B. An evolving hierarchical family classification for glycosyltransferases. *J Mol Biol.* **2003**, *328*, 307.

(94) Lairson, L. L.; Henrissat, B.; Davies, G. J.; Withers, S. G. Glycosyltransferases: structures, functions, and mechanisms. *Annu Rev Biochem.* **2008**, *77*, 521.

(95) Thibodeaux, C. J.; Melancon, C. E.; Liu, H. W. Unusual sugar biosynthesis and natural product glycodiversification. *Nature.* **2007**, *446*, 1008.

(96) Thibodeaux, C. J.; Melancon, C. E., 3rd; Liu, H. W. Natural-product sugar biosynthesis and enzymatic glycodiversification. *Angew Chem Int Ed Engl.* **2008**, *47*, 9814.

(97) Binter, A.; Oberdorfer, G.; Hofzumahaus, S.; Nerstheimer, S.; Altenbacher, G.; Gruber, K.; Macheroux, P. Characterization of the PLP-dependent aminotransferase NikK from Streptomyces tendae and its putative role in nikkomycin biosynthesis. *FEBS J.* **2011**, *278*, 4122.

(98) Burroughs, A. M.; Hoppe, R. W.; Goebel, N. C.; Sayyed, B. H.; Voegtline, T. J.; Schwabacher, A. W.; Zabriskie, T. M.; Silvaggi, N. R. Structural and functional characterization of MppR, an enduracididine biosynthetic enzyme from streptomyces hygroscopicus: functional diversity in the acetoacetate decarboxylase-like superfamily. *Biochemistry.* **2013**, *52*, 4492.

(99) Han, L.; Schwabacher, A. W.; Moran, G. R.; Silvaggi, N. R. Streptomyces wadayamensis MppP Is a Pyridoxal 5'-Phosphate-Dependent l-Arginine alpha-Deaminase, gamma-Hydroxylase in the Enduracididine Biosynthetic Pathway. *Biochemistry.* **2015**, *54*, 7029.

(100) Cai, W.; Goswami, A.; Yang, Z.; Liu, X.; Green, K. D.; Barnard-Britson, S.; Baba, S.; Funabashi, M.; Nonaka, K.; Sunkara, M.; Morris, A. J.; Spork, A. P.; Ducho, C.; Garneau-Tsodikova, S.; Thorson, J. S.; Van Lanen, S. G. The Biosynthesis of Capuramycin-type Antibiotics: IDENTIFICATION OF THE A-102395 BIOSYNTHETIC GENE CLUSTER, MECHANISM OF SELF-RESISTANCE, AND FORMATION OF URIDINE-5'-CARBOXAMIDE. *J Biol Chem.* **2015**, *290*, 13710.

(101) Kaysser, L.; Siebenberg, S.; Kammerer, B.; Gust, B. Analysis of the liposidomycin gene cluster leads to the identification of new caprazamycin derivatives. *Chembiochem.* **2010**, *11*, 191.

(102) Vejpongsa, P.; Yeh, E. T. Prevention of anthracycline-induced cardiotoxicity: challenges and opportunities. *J Am Coll Cardiol.* **2014**, *64*, 938.

(103) Dimarco, A.; Soldati, M.; Fioretti, A.; Dasdia, T. Activity of Daunomycin, a New Antitumor Antibiotic, on Normal and Neoplastic Cells Grown in Vitro. *Cancer Chemother Rep.* **1964**, *38*, 39.

(104) Dimarco, A.; Gaetani, M.; Orezzi, P.; Scarpinato, B. M.; Silvestrini, R.; Soldati, M.; Dasdia, T.; Valentini, L. 'Daunomycin', a New Antibiotic of the Rhodomycin Group. *Nature.* **1964**, *201*, 706.

(105) Dimarco, A.; Gaetani, M.; Dorigotti, L.; Soldati, M.; Bellini, O. Daunomycin: A New Antibiotic with Antitumor Activity. *Cancer Chemother Rep.* **1964**, *38*, 31.

(106) Lown, J. W. The mechanism of action of quinone antibiotics. *Mol Cell Biochem.* **1983**, *55*, 17.

(107) Bonadonna, G.; Monfardini, S.; De Lena, M.; Fossati-Bellani, F. Clinical evaluation of adriamycin, a new antitumour antibiotic. *Br Med J.* **1969**, *3*, 503.

(108) Henry, D. W. Structure-activity relationships among daunorubicin and adriamycin analogs. *Cancer Treat Rep.* **1979**, *63*, 845.

(109) Gewirtz, D. A. A critical evaluation of the mechanisms of action proposed for the antitumor effects of the anthracycline antibiotics adriamycin and daunorubicin. *Biochem Pharmacol.* **1999**, *57*, 727.

(110) Frederick, C. A.; Williams, L. D.; Ughetto, G.; van der Marel, G. A.; van Boom, J. H.; Rich, A.; Wang, A. H. Structural comparison of anticancer drug-DNA complexes: adriamycin and daunomycin. *Biochemistry.* **1990**, *29*, 2538.

(111) Buja, L. M.; Ferrans, V. J.; Mayer, R. J.; Roberts, W. C.; Henderson, E. S. Cardiac ultrastructural changes induced by daunorubicin therapy. *Cancer.* **1973**, *32*, 771.

(112) Smith, B. Damage to the intrinsic cardiac neurones by rubidomycin (daunorubicin). *Br Heart J.* **1969**, *31*, 607.

(113) Tan, C.; Tasaka, H.; Yu, K. P.; Murphy, M. L.; Karnofsky, D. A. Daunomycin, an antitumor antibiotic, in the treatment of neoplastic disease. Clinical evaluation with special reference to childhood leukemia. *Cancer.* **1967**, *20*, 333.

(114) Powis, G. Free radical formation by antitumor quinones. *Free Radic Biol Med.* **1989**, *6*, 63.

(115) Doroshow, J. H.; Davies, K. J. Redox cycling of anthracyclines by cardiac mitochondria. II. Formation of superoxide anion, hydrogen peroxide, and hydroxyl radical. *J Biol Chem.* **1986**, *261*, 3068.

(116) Davies, K. J.; Doroshow, J. H. Redox cycling of anthracyclines by cardiac mitochondria. I. Anthracycline radical formation by NADH dehydrogenase. *J Biol Chem.* **1986**, *261*, 3060.

(117) Kaye, S.; Merry, S. Tumour cell resistance to anthracyclines--a review. *Cancer Chemother Pharmacol.* **1985**, *14*, 96.

(118) Quigley, G. J.; Wang, A. H.; Ughetto, G.; van der Marel, G.; van Boom, J. H.; Rich, A. Molecular structure of an anticancer drug-DNA complex: daunomycin plus d(CpGpTpApCpG). *Proc Natl Acad Sci U S A.* **1980**, *77*, 7204.

(119) Di Marco, A.; Casazza, A. M.; Gambetta, R.; Supino, R.; Zunino, F. Relationship between activity and amino sugar stereochemistry of daunorubicin and adriamycin derivatives. *Cancer Res.* **1976**, *36*, 1962.

(120) Jain, K. K.; Casper, E. S.; Geller, N. L.; Hakes, T. B.; Kaufman, R. J.; Currie, V.; Schwartz, W.; Cassidy, C.; Petroni, G. R.; Young, C. W.; et al. A prospective randomized comparison of epirubicin and doxorubicin in patients with advanced breast cancer. *J Clin Oncol.* **1985**, *3*, 818.

(121) Madduri, K.; Kennedy, J.; Rivola, G.; Inventi-Solari, A.; Filippini, S.; Zanuso, G.; Colombo, A. L.; Gewain, K. M.; Occi, J. L.; MacNeil, D. J.; Hutchinson, C. R. Production of the antitumor drug epirubicin (4'-epidoxorubicin) and its precursor by a genetically engineered strain of Streptomyces peucetius. *Nat Biotechnol.* **1998**, *16*, 69.

(122) Siitonen, V.; Claesson, M.; Patrikainen, P.; Aromaa, M.; Mantsala, P.; Schneider, G.; Metsa-Ketela, M. Identification of late-stage glycosylation steps in the biosynthetic pathway of the anthracycline nogalamycin. *Chembiochem.* **2012**, *13*, 120.

(123) Vejpongsa, P.; Yeh, E. T. Topoisomerase 2beta: a promising molecular target for primary prevention of anthracycline-induced cardiotoxicity. *Clin Pharmacol Ther.* **2014**, *95*, 45.

# Chapter 2: Identification of the Amipurimycin and Miharamycins Gene Clusters Through a Comparative Genomics Approach

## 2.1. INTRODUCTION

Recent advances in the capability to obtain high-quality genomic sequencing data inexpensively through next generation sequencing (NGS) technology have led to the widespread use of genome sequencing as a primary research tool in the study of natural products. Whereas within the last decade the identification and cloning of biosynthetic gene clusters could be a tedious, years-long endeavor, today, identifying natural product biosynthetic gene clusters, from sequence data collection to data processing and annotation, has become a days- or weeks-long project. NGS techniques allow not only very high throughput identification of biosynthetic gene clusters and, often, prediction of gene clusters' products, but also the ability to study natural product gene clusters independent of axenic cultures of the organisms from which they originate.[1-4]

The explosion in the number of available sequenced genomes is not without its challenges. Because many aspects of data collection and processing can be automated, the rate at which newly sequenced genomes are made available has been exponential.[1,2,5] Since the annotation and, hence, prediction of the product of a gene relies on the collected validation information of the function of homologous gene products, the rate of generation of such massive quantities of genomic sequence data greatly outstrips our ability to directly elucidate the function of these genes, and the comparative functional assignment is complicated by mis-annotation of genes in the literature or an inability to annotate genes.[2,3,5] Furthermore, because of the former difficulty in identifying biosynthetic gene clusters, the study of natural product biosynthesis required the collection of copious

biochemical data. These data often yielded critical insight into the biosynthetic pathways in question, and frequently enabled the complete elucidation of a given biosynthetic pathway exclusive of a complete gene cluster sequence.[3] Yet, because the identification of a putative gene cluster now often precedes detailed biochemical data, even with the sequenced gene cluster in hand, the biosynthetic route can still be unclear. The practice of "genome mining," is, thus, a more accurate descriptor of the use of genomic sequence data in the study of natural product biosynthesis.[2,4,6] A key discovery in the biosynthesis of natural products was the observation of their modularity and conservation of their gene clusters, such that structurally similar compounds tend to have similarly composed gene clusters. These conserved gene cluster characteristics are typically shared for a given class of compounds and facilitate the rapid identification of additional gene clusters of class members. This logic relies on the previous successful identification of similar gene clusters, so it is often harder to identify the biosynthetic gene clusters for compounds with poorly conserved biosynthetic gene clusters, such as the peptidyl nucleoside antibiotics (PNAs).[3,7-11]

The PNAs comprise a subset of the nucleoside natural products whose study exemplifies the challenges associated with natural products lacking structural conservation with compounds from other more widely studied groups of natural products.[12] Minimally, the PNAs are composed of a nucleoside with at least one pendant amino acid attached via a peptide bond. However, given the diversity of PNA structures, this group of compounds can be best conceptualized as three components: the nucleobase, the "core" saccharide, and the appended amino acid(s). Frequently, each of these components include structures rarely seen, e.g., hypermodified nucleobases, higher-carbon sugars, and non-proteinogenic amino acids.[9,12-14] Representative PNAs are shown in Figure 2.1.

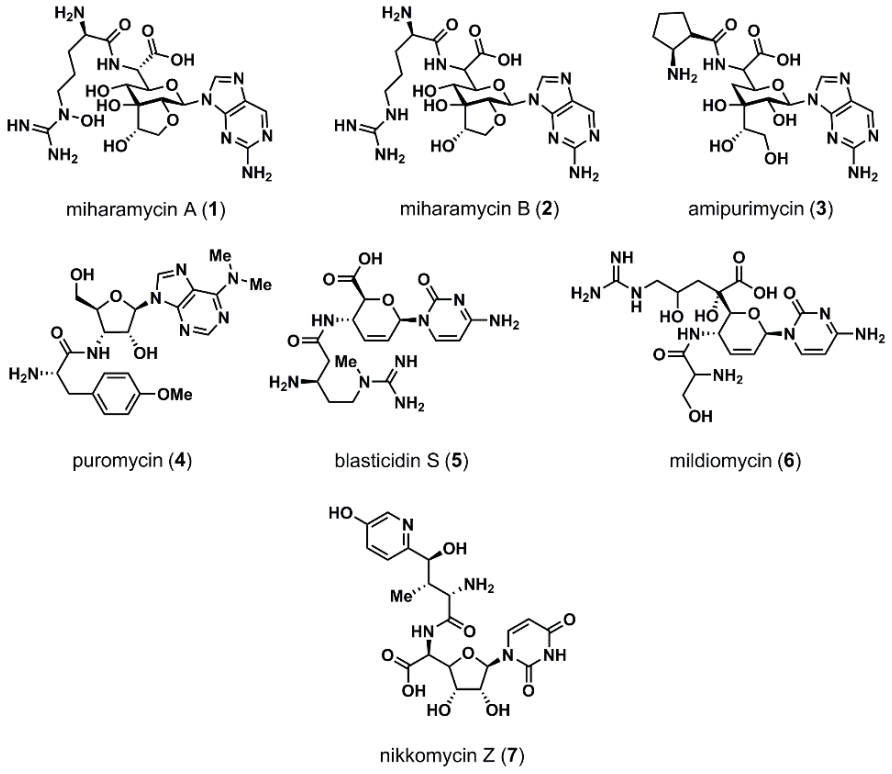



**Figure 2.1:** Structures of representative peptidyl nucleoside antibiotics.

While the biological activities of the majority of the PNAs discovered to date are unknown, several PNAs have been more closely studied, and their activities include: prokaryotic and eukaryotic protein synthesis inhibition,[7,8] prokaryotic tRNA synthetase inhibition,[9] bacterial translocase inhibition,[11,15] and fungal chitin synthase inhibition.[16,17] No PNA are currently clinically used, but nikkomycin Z (**7**) is under investigation as a first-in-class antifungal for the treatment of coccidioidomycosis ("valley fever") in humans.[18-22] Furthermore, puromycin (**4**) and blasticidin S (**5**) are heavily used in molecular biology both as selection agents[23] and probes of ribosomal function;[23,24] and mildiomycin (**6**) finds use as an agricultural fungicide.[19] In light of the diverse biological activities exhibited by these "privileged" pharmacophores and unique constituent amino acids and nucleobases, further study of the PNAs holds promise for both the development

of novel therapeutics and the enrichment of natural product chemistry. Because the structural heterogeneity of PNAs confounds efforts to readily identify their biosynthetic gene clusters, a more complete understanding of PNA biosynthesis is expected to greatly aid in the identification of novel PNA gene clusters.[25,26]

The miharamycins (**1**-**2**) and amipurimycin (**3**) are PNAs elaborated by the soil-dwelling actinomycetes *Streptomyces miharaensis* ATCC 19440 and *Streptomyces novoguineensis* CBS 199.78, respectively, which exemplify the difficulty in finding PNA biosynthetic gene clusters. Despite their discovery over 40 years ago, the structures of these compounds were not elucidated until the 1980s, when they were found to be nucleosides of the rare nucleobase 2-aminopurine.[27,28] Among the PNA, these are the only compounds bearing the 2-aminopurine base. The core saccharides are strikingly similar branched amino nonulosuronic acids, with the fused bicyclic core of the miharamycins resembling a cyclized C4-hydroxy form of the amipurimycin core. The absolute stereochemistry of the molecules was not established in earlier studies, but recently an exhaustive NMR study showed the C6 stereochemistry of miharamycin A is *S*. While this is probably the same for miharamycin B,[29] the C6 stereochemistry of amipurimycin remains undetermined.[27] Similarly, the arginine moiety of miharamycin B was determined to be L-configured; this is probably the case as well for the modified arginine found in miharamycin A, but its configuration has not yet been confirmed.[28] These PNAs also feature the unique amino acids $N^5$-hydroxyarginine and (1*R*,2*S*)-2-aminocyclopentane-1-carboxylic acid, also known as cispentacin (Figure 2.2). Although the biosyntheses of these amino acids have not been studied, they have been isolated from other organisms. $N^5$-hydroxyarginine is known to be produced as a free amino acid by the dermatophytic fungus *Nannizzia gypsea*[30] and two species of Bacillus,[31,32] and is a component of the antibiotics argolaphos A and argolaphos B,[3] as well as asterobactin.[33]

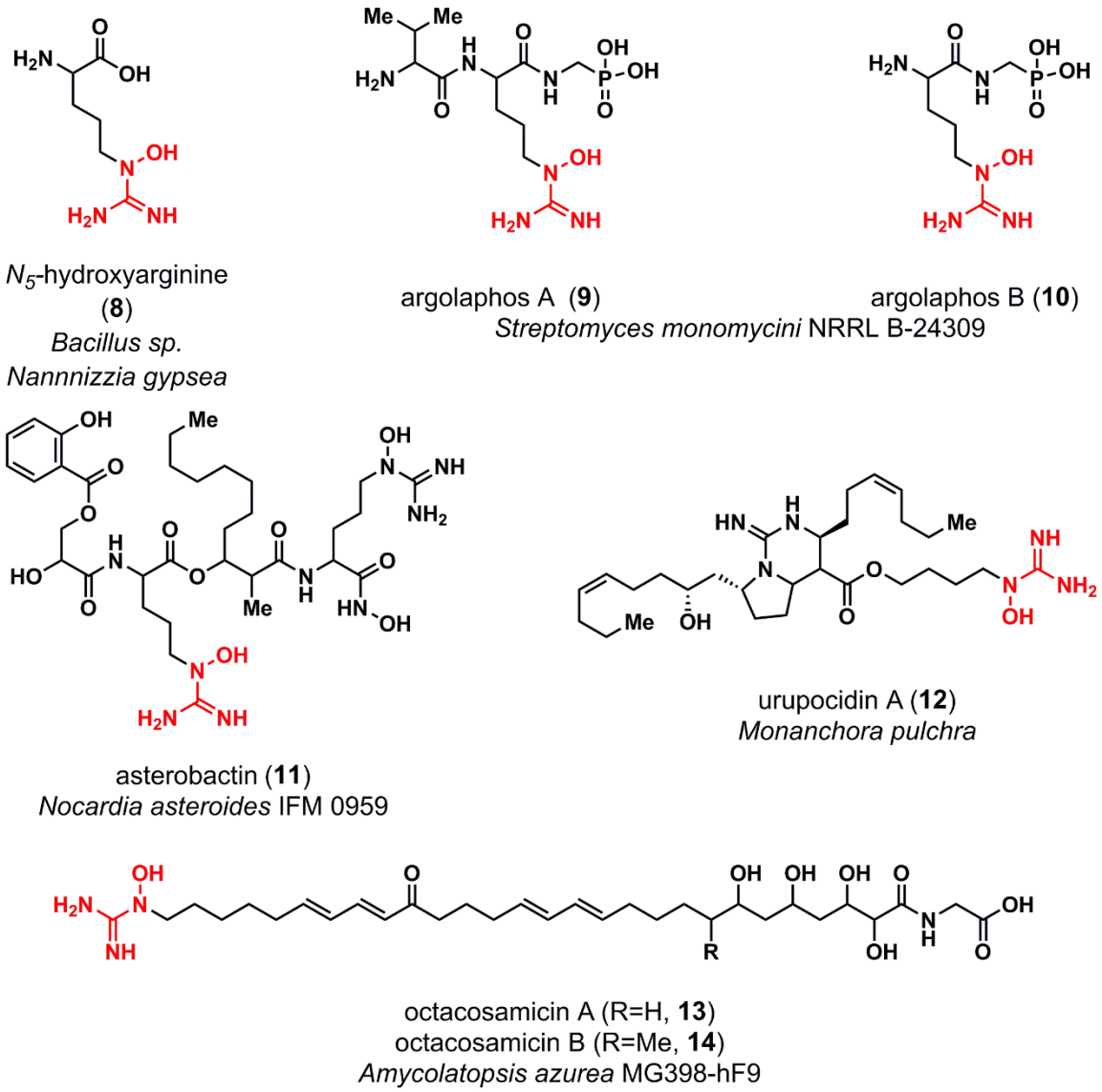



**Figure 2.2:** Biological precedent for $N^5$-hydroxyarginine/ hydroxyguanidine-containing natural products and source organisms.

A similarly hydroxylated guanidine is also found as a component of the marine natural product urupocidin A[34] and the antifungal natural products octacosamicins A and B.[35,36] Cispentacin was found to be produced by a strain of *Bacillus cereus*[37,38] and by *Streptomyces setonii*.[39]

The mechanisms of activity of the miharamycins and amipurimycin are not known, but these PNAs are notable for their potent and specific activity against the phytopathogenic fungus *Magnaporthe grisea* (also known as *Pyricularia oryzae*), the causative agent of rice blast disease,[40,41] as well as activity against a wide variety of phytopathogenic fungi. However, miharamycin A was found to exhibit antiviral activity against tobacco mosaic virus, cucumber mosaic virus, and potato virus X, and the mechanism was suggested to be inhibition of DNA-dependent RNA synthesis.[42,43] The miharamycins are further distinguished from amipurimycin by their strong inhibitory activity in vitro against the Gram negative phytopathogenic bacteria *Pseudomonas aeruginosa*, *Pseudomonas tabaci*, and *Pseudomonas fluorescens*.[40,41,44] Finally, the miharamycins and amipurimycin are considerably toxic to mice when administered intravenously.[40-42]

While the determinants of these biological activities are not known, $N^5$-hydroxyarginine and cispentacin alone exhibit antifungal activity,[30,35,38] and 2-aminopurine has been used as a molecular biology tool for over 50 years in structural studies due to its strong fluorescence and, since it can base-pair with cytosine and thymine, as a bacterial mutagen.[45-48] Despite their potential use in the treatment of rice blast disease and their lack of toxicity to newly emergent leaves, these antibiotics had been shown to damage mature leaves of the plants they were applied to. Nevertheless, the miharamycins and amipurimycin currently have not been used agriculturally or otherwise.

Our interest in the miharamycins and amipurimycin lies in learning the biosynthesis of these compounds' unique structural features, specifically their apparently higher-carbon core saccharides and the rare 2-aminopurine base. The features of these PNA have inspired synthetic efforts over the last almost 30 years, but, surprisingly, neither the miharamycins nor amipurimycin have been successfully totally synthesized.[29,49] In 2008, Marcelo and

co-workers reported the total synthesis of a completely protected form of miharamycin A, but unfortunately attempts to deprotect the molecule resulted in its decomposition.[29] More recently, the same group revealed a novel approach to the bicyclic core of the miharamycins through Wittig olefination of 2-*O*-acyl-D-*ribo*-hexopyran-3-uloses with [(ethoxycarbonyl)methylene]triphenylphosphorane, but the product was still fully protected.[50] Considerably more progress has been made in the synthesis of amipurimycin, as the total synthesis of a thymine analog was reported in 2008 by Stauffer and Datta.[49] Yet, given the rarity of their structures, the miharamycins and amipurimycin remain attractive synthetic targets.

The activity of the miharamycins against *P. aeruginosa*, an important opportunistic human pathogen, is intriguing since the miharamycins do not resemble any antipseudomonal antibiotics in current use.[51,52] Furthermore, the lack of antipseudomonal activity of amipurimycin despite equipotent antifungal activity to the miharamycins suggests distinct structural features underlie these activities. Thus, it is hoped that insight into the biosyntheses of these unique compounds may enable the construction of their synthetically challenging scaffolds which can be used in the development of novel human antimicrobials or chemotherapeutics with improved toxicity profiles. Herein, we describe the identification of the putative amipurimycin and miharamycins biosynthetic gene clusters through a comparative genomics approach. Surprisingly, analysis of these gene clusters suggests the biosyntheses of these compounds rely on enzymes typical of polyketide synthase (PKS) based gene clusters. Furthermore, amipurimycin and miharamycin are proposed to be the first members of a newly discovered group of PNAs.

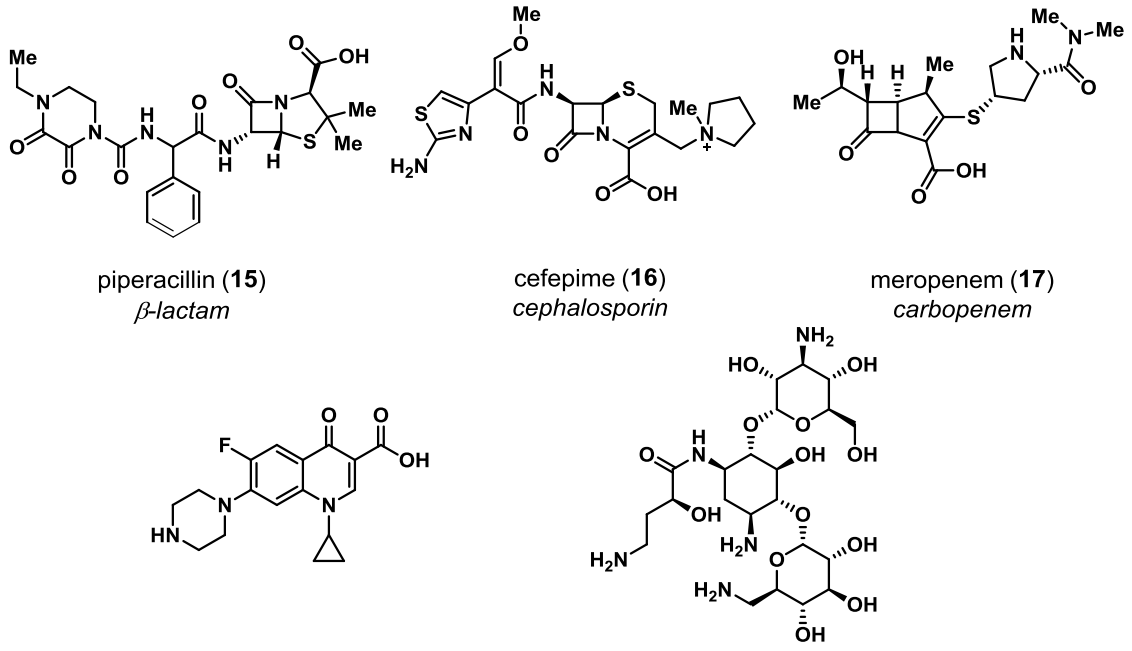



**Figure 2.3:** Antipseudomonal antibiotics in current clinical use and their therapeutic classes.

## 2.2. MATERIALS AND METHODS

### *2.2.1. General*

All chemicals and reagents were purchased from Sigma-Aldrich (St. Louis, MO) or Fisher Scientific (Pittsburg, PA) and used without further purification unless otherwise noted. Laboratory disposables such as microcentrifuge tubes, petri dishes, and conical tubes were obtained from Fisher Scientific or VWR (Radnor, PA). With exceptions as noted, techniques relating to the manipulation of recombinant DNA were performed according to standard protocols;[53] similarly, all *Streptomyces* methods and media were

performed and made according to standard protocols unless otherwise noted.[54] DNeasy Blood and Tissue kits were from QIAgen (Valencia, CA).

### 2.2.2. *Bacterial Strains: Procurement, Preservation, and Growth*

*Streptomyces miharaensis* Niida ATCC 19440 was purchased from American Type and Tissue Culture (Manassas, VA), and *Streptomyces novoguineensis* Iwasa CBS 199.78 was obtained from the CBS-KNAW Fungal Biodiversity Centre (Utrecht, The Netherlands). Both strains were maintained as mycelial suspensions in 20% glycerol and stored at -80 °C in 2 mL Corning® cryo-tubes from Sigma-Aldrich. Strains were routinely grown as 10 mL cultures in sterile seed medium (pH 7.0),[55] composed of 2% glucose, 1% Bacto® peptone from BD (Franklin Lakes, NJ) or peptone from Fisher Scientific, 0.5% meat extract from BD, 0.5% NaCl, and deionized (di) $H_2O$, in test tubes (25 x 150 mm) containing several acid-washed glass beads. Cultures were aseptically inoculated with 0.2 mL mycelial suspension of either S. miharaensis or S. novoguineensis and capped with Kimble KIM-KAP® culture tube caps from Fisher Scientific. These cultures were then grown at 28 °C with rotary shaking at 220 RPM in an Innova® 4200 Incubator-Shaker from New Brunswick Scientific (Edison, NJ) or a MaxQ® 6000 Incubator-Shaker from Thermo Scientific (Waltham, MA).

*2.2.3. Isolation of Genomic DNA*

For the isolation of genomic DNA, 2-3 day old *S. miharaensis* or *S. novoguineensis* liquid cultures grown as above were gently vortexed with a Standard Mini Vortex from VWR, and then the mycelia from 1 mL culture aliquots were collected by centrifugation at 14,000 x *g* in a 5415R refrigerated tabletop microcentrifuge from Eppendorf (Hamburg, Germany). After carefully removing the supernatants by aspiration, genomic DNA was extracted from the mycelial pellets using the QIAgen DNeasy Blood and Tissue kit following the manufacturer's instructions and using the suggested pretreatment for Gram positive bacteria.

*2.2.4. Whole Genome Sequencing*

Samples of *S. miharaensis* and *S. novoguineensis* genomic DNA (~250 ng/µL) in QIAgen buffer AE (10 mM Tris-HCl and 0.5 mM EDTA, pH 9.0) were submitted to the Genomic Sequencing and Analysis Facility (GSAF) of the Institute of Cell and Molecular Biology (ICMB) at the University of Texas at Austin. The genomic DNA libraries for each strain were prepared by the GSAF and sequenced using Illumina MiSeq® technology. Paired-end (PE) and mate-pair (MP) libraries were generated for both strains with target sizes of 2 x 300 $\pm$ 50 base-pairs (bp) and 3-4 kbp, respectively, and 2 million reads were targeted for each strain.

*2.2.5. Data Analysis*

The PE and MP read data for each strain were processed separately using *cutadapt* v. 1.7.1[56] to remove sequencing adapters and isolate reverse-forward reads, respectively. The sequences of the adapters utilized are presented in Table 1. For the PE read data, processing was limited to removal of the Illumina TruSeq® sequencing adapter from the 5'-ends of the PE reads using the following commands:

cutadapt -a TS -o PE_Read1_trim.fastq PE_Read1.fastq

and

cutadapt -a TS -o PE_Read2_trim.fastq PE_Read2.fastq

where the input files were "PE_Read1.fastq" and "PE_Read2.fastq," and the output files were "PE_Read1_trim.fastq" and "PE_Read2_trim.fastq."

| Name | Sequence (5' to 3') | Note | Ref |
|---|---|---|---|
| TruSeq V3 Y adapter ("TS") | GATCGGAAGC | PE Data | 57 |
| Nextera Read 1 External Adapter ("EXT1") | GATCGGAAGAGCACACGTCTGAACTCCAG TCAC | MP Data | 58 |
| Nextera Read 2 External Adapter ("EXT2") | GATCGGAAGAGCGTCGTGTAGGGAAAGA GTGT | MP Data | 58 |
| Circularized Single Junction Adapter ("SJA") | CTGTCTCTTATACACATCT | MP Data | 58 |
| Circularized Single Junction Adapter, Reverse Complement ("rcSJA") | AGATGTGTATAAGAGACAG | MP Data | 58 |

**Table 2.1:** Adapter sequences utilized in whole-genome sequencing.

MP data were processed in four steps. First, reads containing the Illumina Nextera® Mate-Pair adapters EXT1 and EXT2 were removed from the dataset. In order to keep the

read tables aligned, removal of EXT1- and EXT2-containing reads was accomplished using the following command sequence recommended in the cutadapt documentation:

a. cutadapt -b EXT1 -O10 -m100 --discard --paired-output tmp.2.fastq -o tmp.1.fastq MP_Read1.fastq MP_Read2.fastq
b. cutadapt -b EXT2 -O10 -m100 --discard --paired-output Trim_MP_Read1.fastq -o Trim_MP_Read2.fastq tmp.2.fastq tmp.1.fastq
c. rm tmp.1.fastq tmp.2.fastq

where "-O10" specified at least 10 bp matching, "-m100" removed reads shorter than 100 bp, and "MP_Read1.fastq" and "MP_Read2.fastq" were the input files. After removal of the EXT1- and EXT2-containing reads, the remaining reads are represented by the files "Trim_MP_Read1.fastq" and "Trim_MP_Read2.fastq." The last step was used to delete the temporary files "tmp.1.fastq" and "tmp.2.fastq." Next, to obtain reverse-forward oriented MP reads, the reads lacking the Illumina Nextera® Single Junction Adapter (SJA) and its reverse complement (cSJA) were discarded. To keep the two file tables aligned, similarly to the last step, the following command sequence was used:

a. cutadapt -b SJA --trimmed-only --paired-output tmp.2.fastq -o tmp.1.fastq Trim_MP_Read1.fastq Trim_MP_Read2.fastq
b. cutadapt -b cSJA --trimmed-only --paired-output MP_RF_Read1.fastq -o MP_RF_Read2.fastq tmp.2.fastq tmp.1.fastq
c. rm tmp.1.fastq tmp.2.fastq

The final, reverse-forward oriented reads are represented by the files "MP_RF_Read1.fastq" and "MP_RF_Read2.fastq." Finally, the processed MP reads were reverse complemented using the program fastx59with the following commands:

fastx_reverse_complement -i MP_RF_Read1.fastq -o MP_RF_Read1rc.fastq -Q 33

and

fastx_reverse_complement -i MP_RF_Read2.fastq -o MP_RF_Read2rc.fastq -Q 33

The "-Q33" modifier is used to indicate Illumina data are being used since the quality scores from this platform are offset by a value of 33, i.e., quality (Q)= Phred score+33.[60] The final MP read data files ready for assembly are represented by the files "MP_RF_Read1rc.fastq" and "MP_RF_Read2rc.fastq."

*2.2.6. Assembly*

To generate the draft genome sequences for *S. miharaensis* and *S. novoguineensis*, the strains' respective 5'-adapter trimmed-PE reads and reverse complemented, reverse-forward oriented-MP reads were assembled with *velvet*, v. 1.2.10 in two steps, as directed by the manual.[61-63] The dataset for assembly for each strain was created using velveth with following command:

velveth STRAIN_NAMEdir 73 -shortPaired -separate -fastq dir/ PE_Read1_trim.fastq dir/ PE_Read2_trim.fastq -shortPaired2 -separate –fastq dir/MP_RF_Read1rc.fastq dir/MP_RF_Read2rc.fastq

where the kmer length is specified by "73," and the file names are as described above. Once complete, the datasets were stored in directories named for each strain ("STRAIN_NAMEdir"), and assembly was carried out using velvetg with the following command:

velvetg STRAIN_NAMEdir -exp_cov auto -cov_cutoff auto -shortMatePaired yes

where the "-exp_cov" and "-cov_cutoff" parameters were left as their defaults ("auto"), "-shortMatePaired yes" was used to indicate the use of mate-paired data in the dataset, and the final assembly file was saved to the STRAIN_NAME directory.

*2.2.7. Annotation and gene cluster analysis*

The assembled genomic data for each strain in fastq format were uploaded to Rapid Annotation Using Subsystem Technology (RAST) for annotation.[64,65] The annotated genomes were analyzed using RAST tools and antiSMASH 2.0[66] and by whole-genome alignment using *Mauve*.[67-69]

For comparative studies, the draft genome sequence data of *Streptomyces* sp. S-1868, *Streptomyces varsoviensis* NRRL B-3589, *Streptomyces canus* str. 299MFChir4.1, and *Actinoplanes* sp. N902-109, and *Streptomyces monomycini* NRRL B-24309 were

downloaded from NCBI in fastq format via their respective RefSeq FTP sites. These data were uploaded to RAST for annotation and analyzed using the RAST tools.

## 2.3. RESULTS AND DISCUSSION

### *2.3.1. Gene Cluster Identification*

As discussed earlier, the high structural conservation between amipurimycin and the miharamycins, led us to believe they derived from similar or the same components. Thus, the content of their biosynthetic gene clusters was expected to be highly conserved as well. Therefore, comparing the genomes of both producing organisms would allow the identification of conserved gene clusters and facilitate the co-discovery of the amipurimycin and miharamycins gene clusters. The availability of Illumina technology at the GSAF made it possible to quickly sequence with high quality the genomes of both *S. miharaensis* and *S. novoguineensis*, an option previously unavailable to our lab. However, the amount of data this service provided required a strategic approach to analysis.

The applications of next generation genomic sequencing technology are manifold, so unavailability of standard strategies to assemble the data is unsurprising.[70-72] However, given the average size of actinomycete biosynthetic gene clusters of 30-50 kbp,[54] we reasoned identification of the amipurimycin and miharamycins gene clusters would require a final assembly which, even if fragmentary, contained contiguous sequences of at least 50 kbp. Thus, our strategy involved the use of MP read data to generate scaffolds spanning long distances that could be complemented with shorter PE read data.[70-72] To provide a reasonable depth of coverage and, hence, reliability of sequence data, based on a typical

actinomycete genome size of 12 Mbp,[54] we estimated 2 million reads to be sufficient. Processing of the collected data was carried out as summarized in Figure 2.3.

While the removal of the PE sequencing adapter was straightforward, processing of the MP read data to obtain the most useful fraction of the total reads was more difficult. According to the Illumina Nextera® mate-pair library documentation, the MP data are generated from circularized DNA fragments whose ends are linked by the annealing of two junction adapters. These circularized fragments are then digested to provide, besides DNA lacking adapters, the desired reverse-forward (RF) mapping fragments composed of duplex junction adapters flanked by "arms" of DNA. The RF mapping duplexed junction fragments are then fitted with external adapters for sequencing.[58] However, in practice, the MP data contain reads from many other types of fragment-adapter combinations, so it was necessary to: 1) exclude reads containing the external adapter sequence, since these reads are expected to be derived from very short DNA fragments; and 2) keep only reads derived from the ideal DNA flanked duplex junction adapters, since these reads should contain the most useful sequencing data. Finally, due to the RF orientation of the MP read data, these data were reverse complemented.[58]

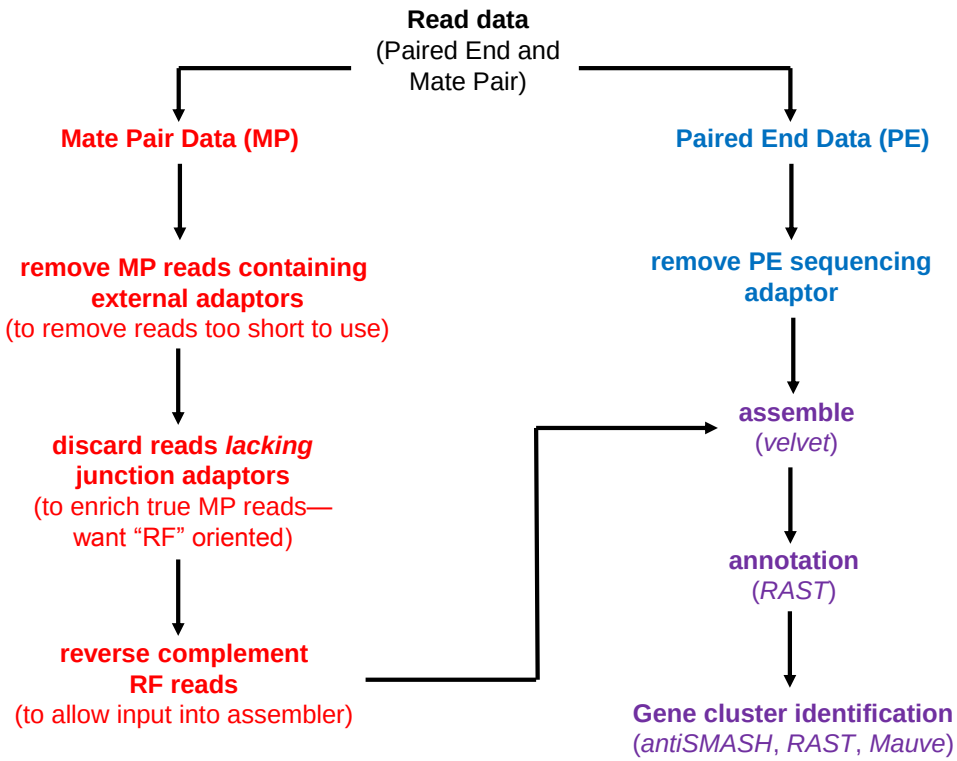



**Figure 2.3:** De novo whole genome sequencing data assembly strategy overview.

Trial assemblies with velveth showed a kmer length of 73 provided a good balance of specificity and coverage,[73] and the best quality assembly was obtained when velvetg was allowed to optimize the parameters "exp_cov" and "cov_cutoff," so they were left as their defaults, which were both "auto." After assembly of the PE and MP read data for each strain, the *S. miharaensis* genome was estimated to be 9.88 Mbp, and the assembly was composed of 223 nodes with an N50 of 1,067824 bp; the *S. novoguineensis* genome was estimated to be 8.32 Mbp, and the assembly was composed of 410 nodes with an N50 of 84,005 bp. After uploading the assemblies to the RAST server, the nodes were converted

to contigs, and 607 contigs were made for *S. novoguineensis*, and 864 contigs were made for *S. miharaensis*. These characteristics indicated sequencing did provide adequate coverage, and there would be a high probability of finding the expected biosynthetic gene clusters on one contig.

The next challenge was the identification of the correct amipurimycin and miharamycins biosynthetic gene clusters. Gene cluster prediction by antiSMASH 2.0 identified over 50 putative biosynthetic gene clusters for each strain, but only one cluster was found to be conserved between both S. miharaensis and S. novoguineensis. Mauve[67] was used to confirm the presence of the shared cluster and localize it to a 30-40 kbp locus in each genome. BLASTP analysis of the predicted products of the predicted ORFs found in these regions using the NCBI Non-Redundant (NR) Protein database suggested these putative clusters were not conserved in all bacteria, further supporting their uniqueness. Annotation by RAST predicted 28 ORFs in the 33.77 kbp conserved region found in *S. novoguineensis* (the putative *snv* cluster) and 27 ORFs in the 30.1 kbp conserved region found in *S. miharaenesis* (the putative *smh* cluster). The clusters are shown in Figure 2.4, and ORFs from both putative gene clusters are listed in Table 2, along with their predicted products based on NCBI BLAST Conserved Domain Database (CDD) and BLASTP analyses. PNAs are conceptually tripartite structures consisting of a core saccharide, a nucleobase, and at least one amino acid.[9,11] Thus, a PNA biosynthetic gene cluster would be expected to encode a means to generate or procure a core saccharide, attach a nucleobase to the core, and attach an amino acid to the core nucleoside. Additionally, PNA clusters

frequently include the enzymes necessary to create any unusual amino acids and nucleobases attached to the core.

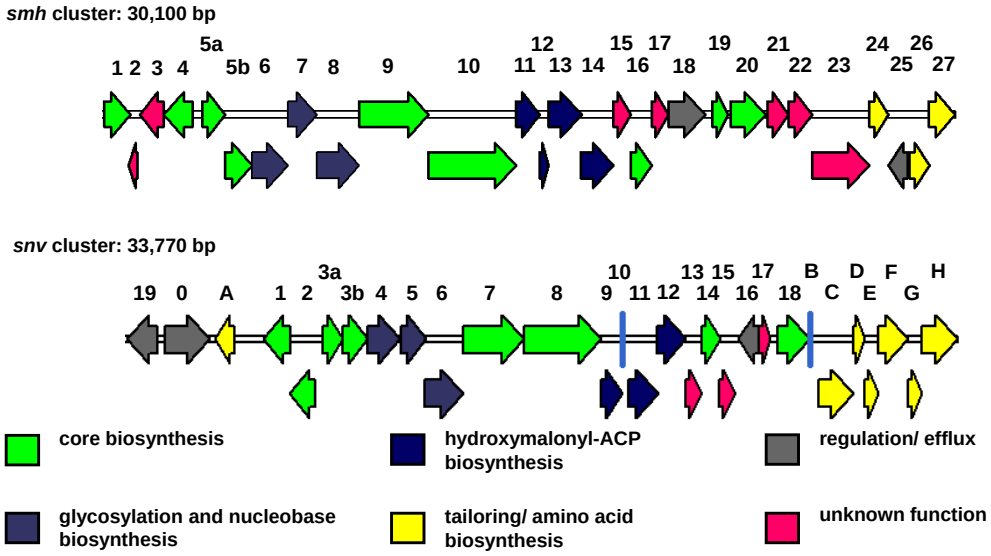



**Figure 2.4:** Organization of the identified putative biosynthetic gene clusters *snv* and *smh*.

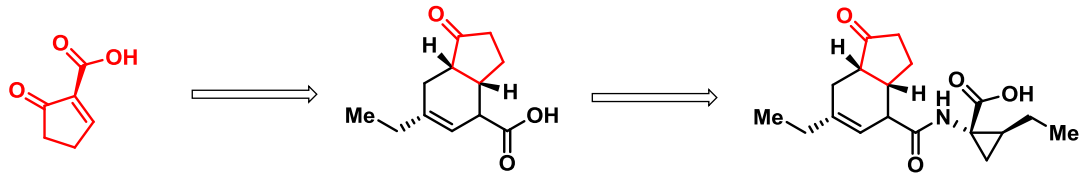
Preliminary investigation of the newly identified *snv* and *smh* gene clusters revealed each encoded enzymes typical of polyketide biosynthesis including a monomodular type I PKS, Snv8 and Smh10, and a subcluster of enzymes implicated in the biosynthesis of the unusual PKS extender unit hydroxymalonylate (as its acyl carrier protein thioester), Snv9-12 and Smh11-14. This finding cast doubt on the role of these clusters in the biosynthesis of amipurimycin and the miharamycins. Subsequent analysis of the *S. novoguineensis* cluster, *snv*, offered support for the role of this gene cluster in the biosynthesis of amipurimycin.

| | | Shared | *S. miharaensis*-specific | | *S. novoguineensis*-specific | |
|---|---|---|---|---|---|---|
| **SMH** | **SNV** | **Proposed Function** | **SMH** | **Proposed Function** | **SNV** | **Proposed Function** |
| 4 | 1 | inositol 2-dehydrogenase | 2 | extradiol dioxygenase α | 0 | pathway activator (SARP) |
| 1 | 2 | branched chain aminotransferase | 3 | extradiol dioxygenase β | A | 3-oxo-ACP reductase |
| 5a | 3a | transketolase N-terminal domain | 19 | PP-transferase ("Sfp") | B | ACP |
| 5b | 3b | transketolase C-terminal domain | 21 | UbiE methyltransferase | C | Orn aminotransferase |
| 6 | 4 | CGA synthase homolog | 22 | Phe-tRNA synthetase α | D | thioesterase |
| 7 | 5 | aminoglycoside kinase | 23 | Phe-tRNA synthetase β | E | 3-oxo-ACP dehydr. |
| 8 | 6 | flavin-dept. oxidoreductase | 24 | YqcI/YcgG | F | 3-oxoacyl-ACP synth. |
| 9 | 7 | PKS loading module, A-T-type | 25 | TetR regulator | G | Dieckmann cyclase |
| 10 | 8 | PKS, type I (iterative?), AT-KS-T | 26 | aldolase | H | acyl-CoA ligase |
| 11 | 9 | 3-hydroxybutyryl-CoA dehydr. | 27 | ribose kinase | 16 | pathway activator (SARP) |
| 12 | 10 | ACP | | | | |
| 13 | 11 | short chain acyl-CoA dehydr. | | | | |
| 14 | 12 | FkbH-like | | | | |
| 15 | 13 | UbiE methyltransferase | | | | |
| 16 | 14 | gluconate 5-reductase | | | | |
| 17 | 15 | LOG-1 homologue | | | | |
| 20 | 18 | ATP-grasp ligase | | | | |
| 18 | 19 | MFS-1 efflux protein | | | | |

**Table 2.2:** ORFs in the *snv* and *smh* gene clusters and their predicted products.

We found the majority of the *snv* ORFs appeared to be inserted into a small gene cluster encoding homologs of Cfa1-5, comprising the biosynthetic pathway of 2-carboxy-2-cyclopentanone (CPC). The polyketide-derived compound CPC is a precursor of coronafacic acid in the biosynthesis of coronatine, a virulence factor produced by several

pathovars of *Pseudomonas syringae*.[74-76] The remarkable similarity between CPC and cispentacin suggested the latter could be derived from reduction and transamination of the former. Indeed, an enoyl reductase, SnvA, and an ornithine aminotransferase, SnvC, not found in any other annotated CPC gene clusters in the NCBI database were encoded in the same genetic locus, which provided the first support for the correct identification of the amipurimycin gene cluster. This pathway will be discussed in depth in a later section.




2-carboxy-2-cyclopentanone (**20**) coronafacic acid (**21**) coronatine (**22**)

**Figure 2.5:** Structures of coronatine and intermediates.

To ensure the completeness of the snv and smh clusters, we performed comparison with other genomes in the RAST database in an effort to delineate the snv and smh gene cluster boundaries. Surprisingly, the snv cluster appears to lie within a conserved locus in *Streptomyces coelicolor A3(2)*[77] corresponding to *sco6992* to *sco7000* which supports the identified cluster boundaries. The *smh* cluster boundaries are not as obvious upon comparison with other genome sequences in the RAST database, but its high conservation to the *snv* cluster supports the identified cluster boundaries.

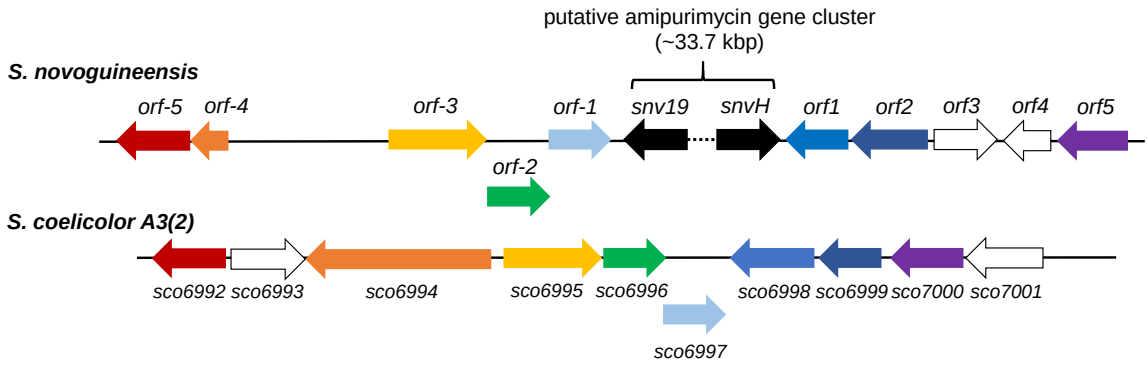



**Figure 2.6:** Proposed *snv* gene cluster boundaries deduced from comparison with the *S. coelicolor A3(2)* genome.

*2.3.2. Gene Cluster Analysis*

Overall, the clusters are notable for their organization as three shared major portions: a three gene cassette putatively encoding a glycosyltransferase (Snv4 and Smh6), an aminoglycoside resistance kinase (Snv5 and Smh7), and flavin-dependent oxidoreductase of unknown function (Snv6 and Smh7); a monomodular minimal type I polyketide synthase (PKS) composed of single ketosynthase (KS), acetyltransferase (AT) and phosphopantetheinyl binding/ thiolation (T) domains immediately preceded an unusual putative adenylation/thiolation didomain loading module (Snv7 and Snv8; Smh9 and Smh10); and a conserved subcluster for the biosynthesis of hydroxymalonyl-ACP (Snv9-12 and Smh11-14). An ATP-grasp ligase (Snv18 and Smh20), a branched-chain aminotransferase (BCAT; Snv2 and Smh1), a two-component transketolase (Snv3a/b and Smh5a/b), a UbiE methyltransferase (Snv13 and Smh15), a putative LOG-1 homolog (Snv15 and Smh17), and two oxidoreductases, an inositol 2-dehydrogenase (Snv1 and Smh4) and a gluconate 5-reductase (Snv14 and Smh16), and a putative resistance efflux protein (Snv19 and Smh18) are also shared.

Inspection of the putative biosynthetic gene clusters identified in S. miharaensis and S. novoguineensis revealed their content varied from any previously published PNA gene clusters, making comparison difficult. However, the smh and snv clusters do contain a weak CGA synthase homolog (Smh6, Snv4), a hallmark of several cytosyl-PNA gene clusters, but they lack the nucleoside hydrolase found in the known cytosyl-PNA gene clusters.[9] This is not unexpected, since unlike these PNAs, the miharamycins and amipurimycin bear a purine base instead of cytosine, which is found in very low abundance as a free base physiologically.[78] Nevertheless, the presence of the glycosyltransferases in the snv and smh clusters suggests, as is the case for the cytosyl-PNAs,[9] amipurimycin and the miharamycins are derived from glycosylation of a nucleobase with a sugar and not from modifications to a preformed nucleoside or nucleotide, a strategy employed in the biosyntheses of some PNAs, such as puromycin,[24] the nikkomycins,[17] and the pacidamycins.[79]

*2.3.2.1. Core Saccharide Biosynthesis*

*2.3.2.1.1. Involvement of PKS enzymes*

Based on the biosynthetic gene clusters of blasticidin S,[23] mildiomycin,[80] and the other cytosyl-PNAs,[81,82] one pathway envisioned for the biosynthesis of amipurimycin and the miharamycins involves the glycosylation of a purine base with an NDP-activated hexose, followed by elaboration via the PKS/ transketolase and amino acid attachment by the ATP-grasp ligase, but the origin of the requisite NDP-sugar is not clear. The nine carbon core saccharides of the miharamycins and amipurimycin are rare amongst the PNAs, as the highest carbon sugars identified are sedoheptoloses, such as spicamycin and septacidin[83] or octuloses, like ezomycin A1.[13,14] More importantly, the miharamycins and

amipurimycin include both amino and carboxylate functionalities, a characteristic shared only by a handful of PNAs,[16,17,84,85] as well as some liponucleosides.[86-88] With the exception of the octosyl acid derivatives, the dual amino and carboxylate functionalities in these compounds arise from condensation of a nucleoside with glycine though the action of a serine hydroxymethyltransferase (SHMT) which functions as a pyridoxal monophosphate (PLP)-dependent aldolase.[83] Octosyl acid-derived PNAs are generated from the condensation of uridine-5'-monophosphate and phosphoenol pyruvate, and then this product can be subject to tailoring reactions such as carbon atom elimination and transamination, albeit via poorly characterized mechanisms.[13,14,16] Because the *snv* and *smh* gene clusters contain neither SHMTs nor enzymes indicative of the intermediacy of octosyl acid, we considered the role of the PKS enzymes encoded in these cluster.

Intriguingly, the clusters are dominated by enzymes characteristic of polyketide natural product biosynthetic gene clusters, raising the possibility the branched core "saccharides" of amipurimycin and the miharamycins could be functionalized polyketide products. However, there is no precedent for the formation of a sugar by a PKS *per se*, and the potential involvement of such an enzyme is complicated by the uncommonly simple architecture of the PKS enzymes in these clusters, as well as the branched

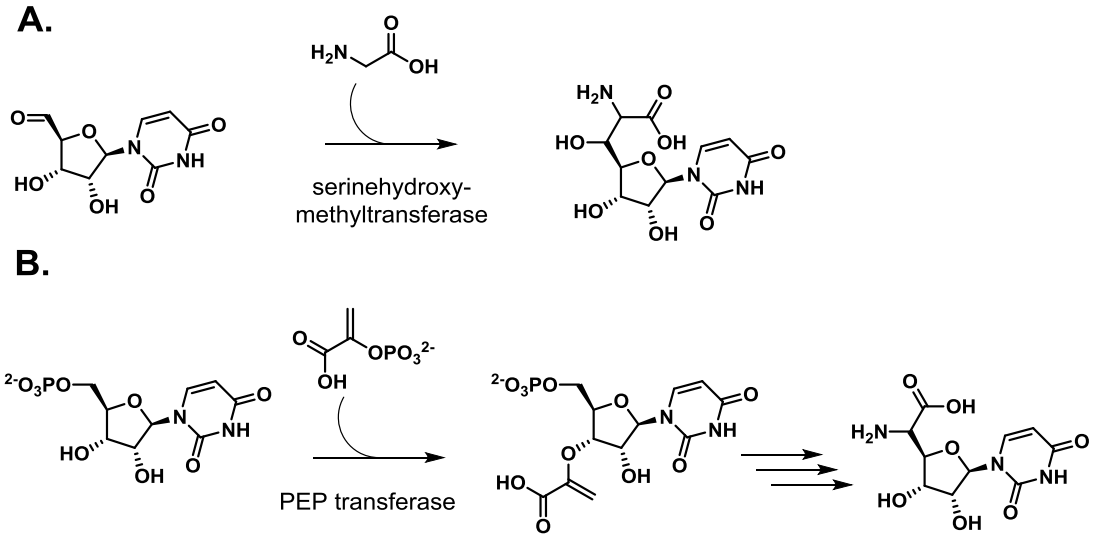



**Figure 2.7:** Known methods to install dual amino and carboxylic acid functionalities in PNAs and related natural products.

structure of the saccharides, since, according to standard PKS logic, the nascent products of such a basic PKS system would be linear.[7,8] Typically, type I PKS enzymes are very large and composed of several modules. Each module is minimally composed of ketosynthase (KS), acyltransferase (AT), and thiolation (T) domains, and each module is responsible for activating malonyl- or acetyl-CoA for the extension of a starter unit supplied by the initial or loading module and then the product of the preceding module by two carbon units.[89-91] Although several mechanisms for PKS product release have been observed and proposed, the nascent polyketide chain is then typically released by a thioesterase (TE) domain usually found situated in the terminal module, and cyclization frequently ensues, generating a lactone product.[92] Therefore, the length of a polyketide product can generally be predicted based on the number of modules composing the PKS.[7,8]

Following this biosynthetic logic and since the core saccharides of the miharamycins and amipurimycin are branched and contain nine carbon atoms, the PKS should contain at least a loading module, more than one extension module, and some means of branching. Furthermore, because each extension module adds only two carbon atoms, an unusual three- or five-carbon starter unit would be necessary. However, as mentioned earlier, the PKS enzyme found in the miharamycins and amipurimycin biosynthetic gene clusters is composed of a single minimal module lacking a TE, and, in contrast to the PKSs involved in the biosynthesis of branched polyketides, no PKS branching domains or other previously described branching enzymes are present.[93-96] Curiously, the smh and snv clusters each include a two-component transketolase encoded by Smh5a/5b and Snv3a/3b, respectively, which could be utilized to effect branching as a post-PKS tailoring reaction.

Some of these peculiarities can be reconciled by recent examples of deviations from standard PKS logic.[97,98] Most importantly, it is now well-known that the co-linearity suggested by the multi-modular arrangement of PKS enzymes cannot be taken as a given. Frequently, one or more modules are used more than once in the generation of a linear polyketide. This so-called "stuttering" activity is difficult to predict a priori and is usually only apparent upon inspection of the final product.[7,8,99] Thus, it is not unreasonable to propose the PKS enzymes found in the SMH and SNV clusters are used iteratively. Although originally thought to be limited to fungi, iterative type I PKSs are also found in the biosynthetic gene clusters of several aromatic bacterial metabolites. However, recently several examples of bacterial iterative PKSs involved in the biosynthesis of non-aromatic reduced metabolites were described. Notably, the enediynes utilize an unusually composed

monomodular type I PKS lacking an obvious C domain to generate an octaketide precursor to their 9- or 10-membered aceytlenic "warhead" portions.[98]

Also, while malonyl-CoA and acetyl-CoA are the most common PKS starter and extender units, the biosynthesis of polyketides can employ a wide spectrum of unusual "ketide" units.[89,90] Hydroxymalonate and methoxymalonate are unique among the alternative ketide units as PKS biosynthetic pathways that utilize these units include a biosynthetic subcluster of a conserved set of four genes to generate hydroxymalonyl-ACP or methoxymalonyl-ACP from 1,3-bisphosphoglycerate.[100-102]

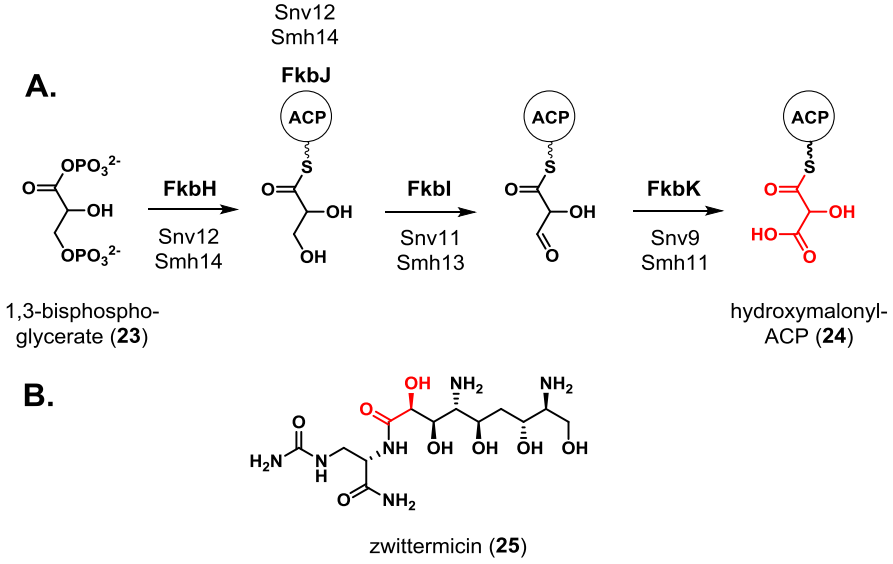



**Figure 2.8:** Methoxymalonyl-ACP/ hydroxymalonyl-ACP biosynthesis via the "Fkb" subcluster and occurrence of hydroxymalonate in zwittermicin.

First found in the FK506 gene cluster, the so-called Fkb subcluster includes FkbH, which encodes an unusual phosphatase responsible for loading 1,3-bisphosphoglycerate onto a specific ACP encoded by FkbJ, and FkbK and FkbI, which successively oxidize the

free primary alcohol to a carboxylic acid to give ACP-linked hydroxymalonate. The cluster can also include an *O*-methyltransferase, FkbG, which methylates the secondary alcohol of hydroxymalonyl-ACP to yield methoxymalonyl-ACP. A notable Fkb-like subcluster lacking an FkbG homolog is found in the *zma* gene cluster involved in the biosynthesis of zwittermicin.[1,2]

Strikingly, the *smh* and *snv* clusters each have an Fkb subcluster found immediately downstream of the PKS-encoding gene, suggesting hydroxymalonyl-ACP may serve as a potential PKS starter and/ or extender unit. The potential use of an uncommon starter unit is supported by the putative alternative didomain adenylation/ thiolase-type PKS loading module (Snv7 and Smh9) encoded immediately upstream of the PKS-encoding ORFs in each cluster. These modules have been recognized to load "unusual" PKS starter units, and, in contrast to other types of loading modules, appear able accept free acid substrates.[3] The exact role the monomodular PKS enzymes found in the *snv* and *smh* gene clusters play in the formation of these pyranoses remains to be established, but the substantial deviations from canonical PKS enzymology described above provide the basis for a biosynthetic model.

*2.3.2.1.2. PKS-centric Biosynthetic Model*

The presence of PKS enzymes in the *snv* and *smh* clusters can be rationalized by a biosynthetic model in which a rare three-carbon starter unit is iteratively elongated, and, as suggested by the Fkb cassettes in these gene clusters, one or more molecules of hydroxymalonyl-ACP are used in the biosynthesis of amipurimycin and the miharamycins. However, a key difference between the core saccharides of the miharamycins and amipurimycin is the C4'-hydroxyl in the former and its absence in the latter. Deoxygenation is a sugar modification that has been observed to require a PLP-dependent

aminotransferase-like enzyme, but no such enzymes are found in the *snv* and *smh* gene clusters besides the single branched-chain aminotransferase these clusters share. Alternatively, we propose that the miharamycins and amipurimycin share hydroxymalonate as a PKS starter unit, but the former are further composed of at least two additional hydroxymalonyl ketide units, whereas the latter is further composed of one additional hydroxymalonyl ketide unit and one malonyl ketide unit.

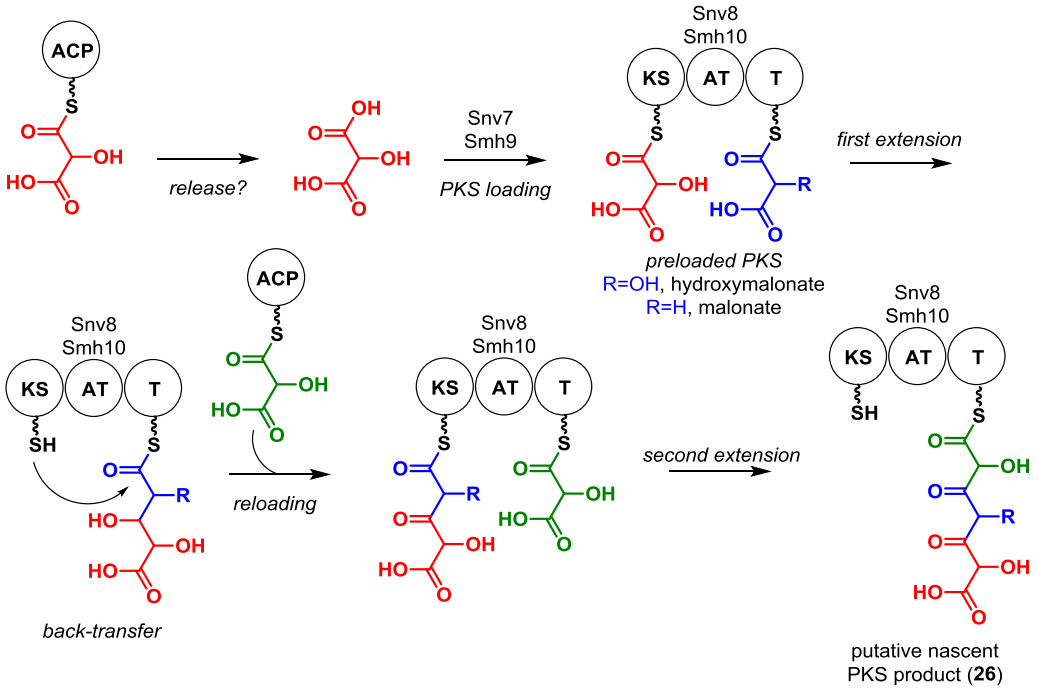



**Figure 2.9:** PKS-centric model for the biosynthesis of the miharamycins and amipurimycin.

Thus, biosynthesis of the miharamycins and amipurimycin commences when Smh9 or Snv7 loads the starter unit to the KS domain of Smh10 or Snv8, respectively. Because hydroxymalonate is produced as its ACP-thioester, its use by an adenylating PKS loading module would require some means to release the free acid from the ACP. A suitable

hydrolase or thioesterase is not immediately obvious based on analyses of the snv and smh gene clusters' content, but both clusters encode a small enzyme, Snv15 and Smh17, respectively, annotated as a "lysine decarboxylase." by NCBI. The basis for this predicted function is not clear because Snv15 and Smh17 are much smaller than authentic lysine decarboxylases. However, these enzymes do have putative Rossmann motifs, indicating binding of a cofactor such as NAD(P)H, and may therefore possess latent function as oxidoreductases.

Homologues of Snv15 and Smh17 have been implicated as molybdenum cofactor (MoCo) carrier proteins[105] or as "Lonely Guy-1" (LOG-1) proteins involved in the maturation of cytokinin plant hormones.[106,107] The function of Snv15 and Smh17 as MoCo loading proteins is doubtful given the absence of MoCo-utilizing enzymes in their respective clusters. On the other hand, LOG-1 proteins putatively function as riboside 5'-phosphoribohydrolases, hydrolyzing cytokinin nucleotides to yield mature cytokinin free bases and ribose 5'-monophosphate. In the latter capacity, it is plausible Snv15 and Smh17 could act to release hydroxymalonate from its ACP as shown in Figure 2.8, either hydrolytically or, in consideration of their predicted Rossmann motifs, perhaps as a reductase type of thioesterase.[92] As is common for type I PKSs, Snv8 and Smh10 are expected to harbor pre-loaded T-domains.[7] In the case of the miharamycins, the pre-loaded ketide unit is predicted to be hydroxymalonate, and to account for the C4-deoxy pyranose core of amipurimycin, as discussed earlier, we believe the T-domain of Snv 8 could be pre-loaded with malonate. After the initial condensation, the five-carbon product is transferred back to the KS domain for one more round of elongation by another molecule of hydroxymalonate to yield a seven carbon linear triketide.

The ensuing release and cyclization steps are unclear, since Smh10 and Snv8 lack a TE domain, and no stand-alone TE is encoded elsewhere in the clusters. Furthermore,

no other means of product release, such as a PKS reductase or cyclase domains, can be readily identified in the *smh* and *snv* clusters,[92] although it is possible this is a role fulfilled by the putative hydrolases Snv15 and Smh17. In vitro experiments using cloned truncated PKS enzymes or individual modules demonstrate spontaneous cyclization and off-loading of nascent products is facile, but the full-length PKS enzymes they derive from usually rely on a distinct thioesterase, whether as a PKS domain or a stand-alone enzyme, for off-loading and cyclization *in vivo*.[8] Alternatively, while no pyran cyclase/ cyclase domains can be identified in Smh10 or Snv8, it could be argued that the product of the *smh* and *snv* PKSs are branched pyrans hydrolytically released. Several mechanisms for pyran formation have been reported, and though the cyclization is effected through different modalities, the conserved mechanism is an oxa-1,4-conjugate addition that occurs on a PKS-tethered intermediate and requires an unsaturated substrate.[93-96] The strongest argument against this possibility is the lack of a mechanism to desaturate the requisite substrate, such as a dehydratase or dehydratase PKS domain.

However, both clusters encode methyltransferases downstream from their Fkb subclusters (Smh15 and Snv13) distinct from FkbG, and whose closest annotated homologs are each other. The function of these enzymes is not clear because observation of the structures of amipurimycin and the miharamycins does not reveal evidence of methylation, but these methyltransferases could methylate the carboxylate of their respective free PKS products, encouraging cyclization. There are no examples of this mechanism from PKS biosynthesis, but in biosynthesis of the capuramycins, methylation is used to activate a carboxylate for amide bond formation, resulting in expulsion of methanol.[108,109] The possibility these methyltransferases function as unusual pyran forming cyclases cannot be ruled out at this point. By analogy to other full-length PKS and NRPS pathways lacking an obvious means to off-load products,[92] and in consideration of the above, we propose

cryptic product release and cyclization by one of at least three sequences (Figure 2.10): a) the triketide product is spontaneously tautomerized while still tethered to the PKS, enabling lactonization and product release; or b) the PKS-bound product is cyclized by a *trans*-acting cylclase, leading to product release; or c.) the PKS product is hydrolytically released, and subsequent tautomerization enables spontaneous or enzyme-catalyzed lactonization. Importantly, tautomerization of the PKS product yields a carbonyl at C6 which is expected to be aminated by the BCAT homologs Snv2 and Smh1 to allow attachment of the amino acids found in amipurimycin and the miharamycins by their respective ATP-grasp ligases, Snv18 and Smh20.

As mentioned earlier, the smh and snv clusters each encode a two component transketolase (Snv3a/3b and Smh5a/5b). In light of the branched core saccharides of the miharamycins and amipurimycin, two roles for the transketolases can be envisioned: the transketolases could interact with the PKS enzymes to generate or load an unusual branched PKS extender unit; or the transketolases are responsible for the addition of the two carbon branch clearly observed in the structure of amipurimycin and, in cyclized form, probably the miharamycins. The juxtaposition of PKS machinery and transketolases is apparently rare, as only two examples of the former case have been reported.

In the first example, the two component transketolase Azi C5/C6 is putatively involved in the formation of a unique aziridine-containing amino acid in the biosynthetic pathway of azinomycin B, an NRPS-derived natural product.[110] In the other example, the products of a conserved three-gene cassette including an unusual transketolase were shown to procure the precursor to a two-carbon branch found in the tetrahydroisoquinoline antibiotics quinocarcin (QNC cluster) and ecteinascidin 743 (NDM cluster).[111] Similarly to Smh5a/b and Snv3a/b, the TPP-binding domain of the transketolases in these clusters is divided between NapB and NapD or Qnc N and QncL, but NapD and Qnc N feature an

additional C-terminal domain reminiscent of pyruvate dehydrogenase component E2. These transketolases were shown to transfer a two-carbon unit from ketose phosphates to the third products of the cassette, the ACPs Nap C and

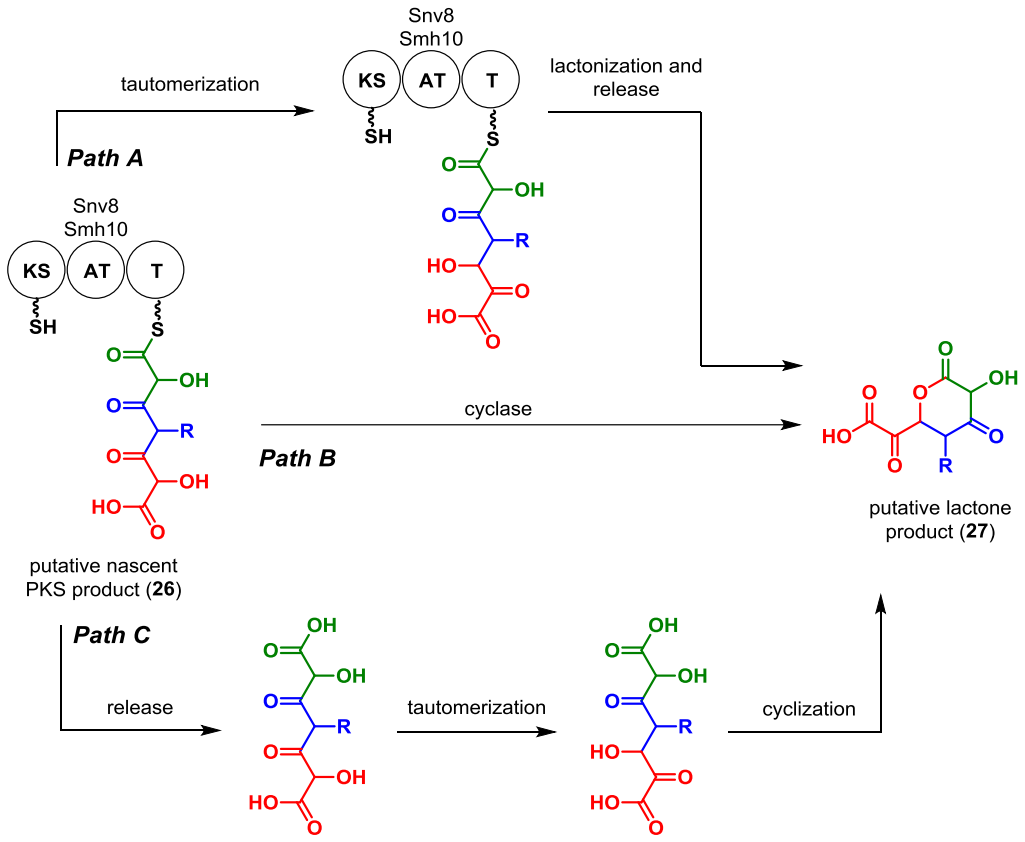



**Figure 2.10:** Hypothetical PKS release mechanisms in the biosynthesis of the miharamycins and amipurimycin.

Qnc M. Once acylated, Nap C and Qnc M interact with the respective NRPS machinery to deliver the two-carbon unit. Given the lack of an E2 domain in the transketolases found in the snv and smh gene clusters, they are not expected to act on a PKS/ ACP-associated substrate.

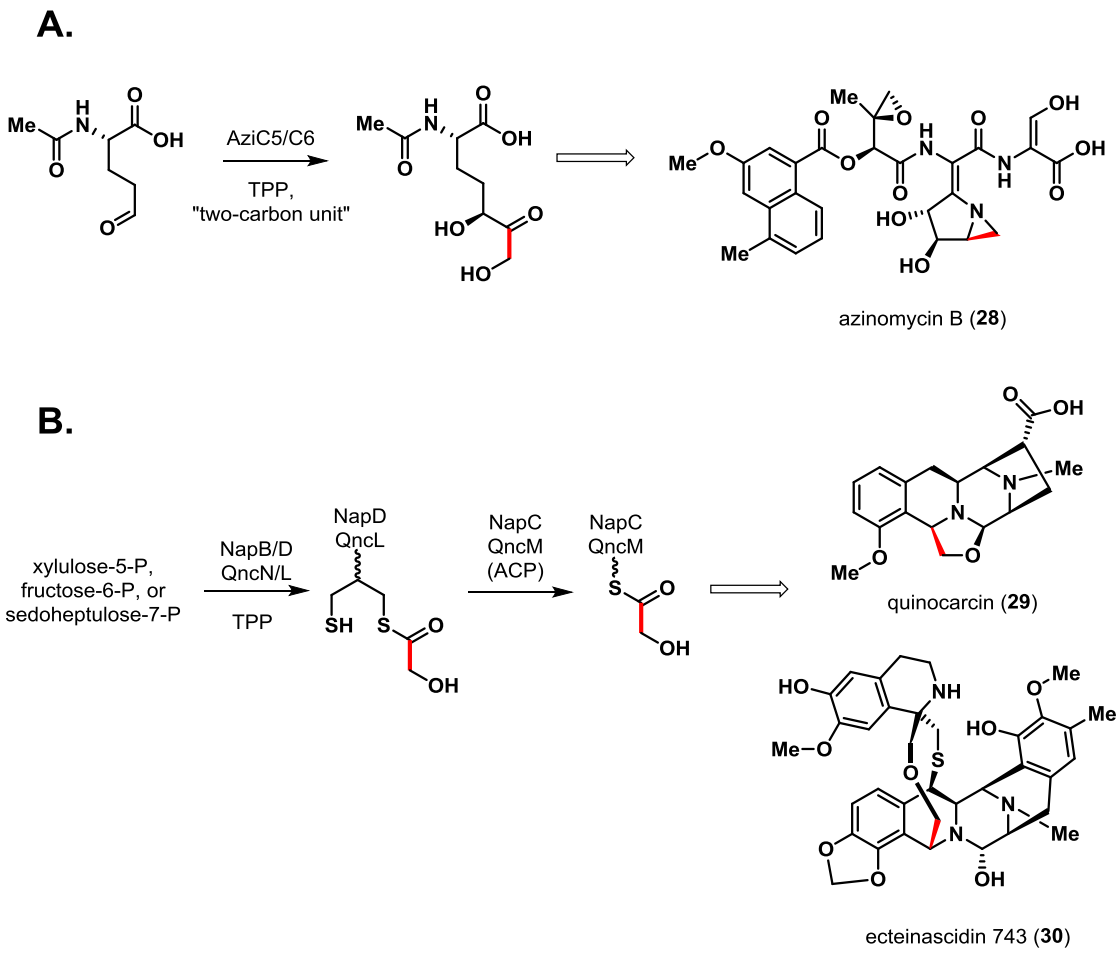



**Figure 2.11:** Examples of thiamine-dependent enzymes in PKS biosynthetic pathways.

As mentioned earlier, two-carbon branching is a well-known tailoring reaction in deoxy sugar biosynthetic pathways.[112] The archetypical transketolase is found in the pentose phosphate pathway, where this TPP-dependent enzyme acts in the interconversion of ketose phosphates. However, transketolase-like enzymes are found in the biosynthetic gene clusters of several natural products containing two-carbon branched deoxy sugars.[83] By analogy to the biosyntheses of hydroxyethyl-branched sugars, we predict Smh5a/b and Snv3a/b, the transketolases found in the snv and smh gene clusters, contribute a two-carbon

fragment to the core heterocycle of the miharamycins and amipurimycin, generating the branched nine-carbon scaffold of these compounds.

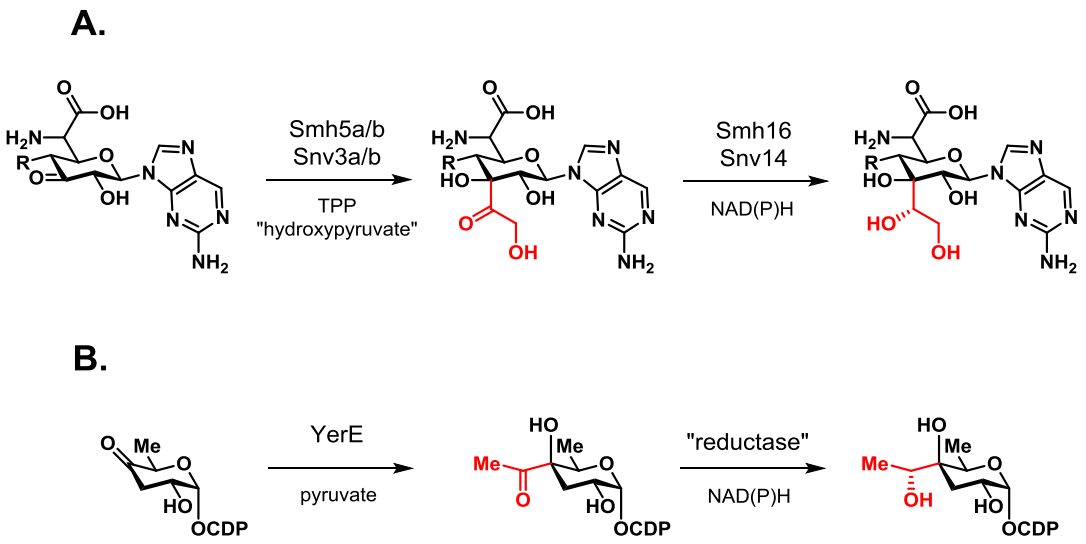



**Figure 2.12:** Transketolase-mediated two-carbon branching and subsequent reduction in the biosynthesis of the miharamycins and amipurimycin compared to a similar pathway from yersiniose biosynthesis.

Although these sugars usually feature α-oxoethyl or α-hydroxyethyl branches derived from pyruvate,[83] we felt the 1,2-dihydroxyethyl branch could probably derive from a reduced glycoaldehyde shunted from primary metabolism. The proposed reduction of the Snv3a/b or Smh5a/b-derived two-carbon branch has precedent in the biosynthesis of CDP-yersiniose in *Yersinia pseudotuberculosis VI*.[113] As encountered in nature, yersiniose features a hydroxyethyl branch, yet the pyruvate-derived two-carbon unit installed by YerE features a carbonyl. While the required reductase was not identified, the yersiniose biosynthetic gene cluster also encodes a putative short-chain dehydratase/reductase (SDR).[113] This arrangement is mirrored in the *snv* and *smh* clusters, which contain the gluconate 5-reductase (SDR) homologues Snv14 and Smh16. Thus, we predict the

installation of the two-carbon branch to be a tailoring step shared in the amipurimycin and miharamycins biosynthetic pathways. After this branch is installed, the reductases Snv14 or Smh16 may act to generate the observed 1,2-dihydroxyethyl appendage found in open-chain form in amipurimycin and, likely, in closed-chain form in the miharamycins.

*2.3.2.1.3. Perhydrofuropyran Formation*

The bicyclic perhydrofuropyran core of the miharamycins is not well represented in natural products. While octosyl acid[16] and the structurally similar core saccharides of the ezomycins and malayamycin,[114-116] which appear to be octosyl acid derivatives, are also perhydrofuropyrans, the ether oxygens of their structures are on opposite sides of the ring system, as opposed to the analogous oxygen atoms on the miharamycin core, which are found on the same side of the ring system, indicating a likely divergent biosynthetic pathway from octosyl acid and the like.

Furthermore, the nucleobase of the octosyl acids is found attached to the furan ring, unlike the miharamycins, where the nucleobase is found on the pyran portion. Indeed, only the biosynthesis of octosyl acid has been studied, and this compound was found to be derived from the condensation of phosphoenolpyruvate and uridine 5'-monophosphate by the action of a unique transferase.[13,14,16,117] While not a PNA, the natural product SB-219383 is a bicyclic iminosugar which bears some similarity to the core saccharide of the miharamycins, but its biosynthetic gene cluster has not been identified.[118]

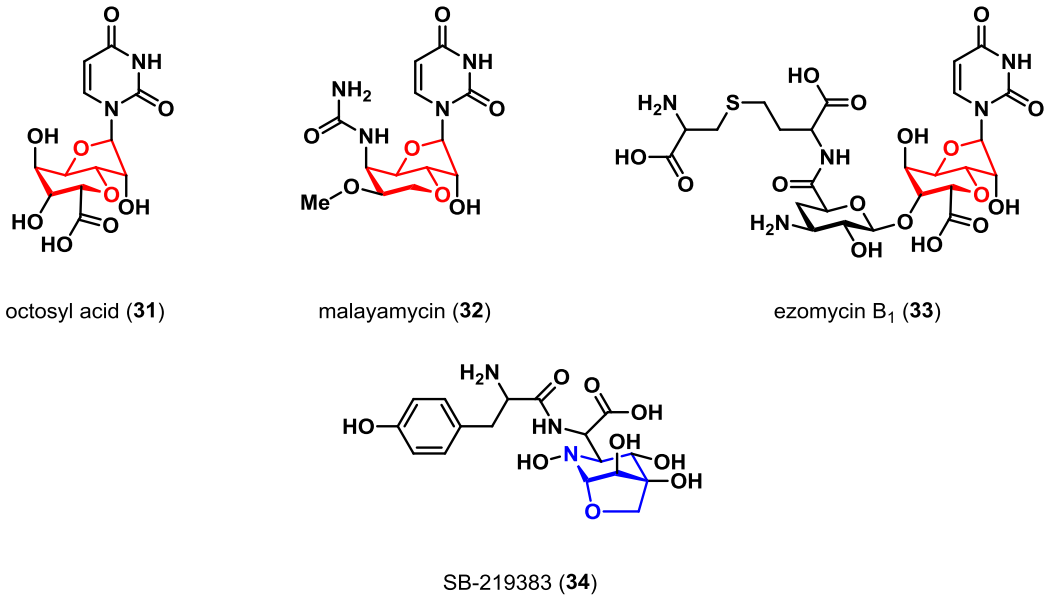



**Figure 2.13:** Natural occurrence of the perhydrofuropyran structural motif and a similar [3.2.1] aza-bicycle.

Early in our studies, we believed the perhydrofuropyran of the miharamycins could be formed similarly to octosyl acid, but the content of the newly identified smh gene cluster does not support this pathway. Instead, it was assumed the unique bicyclic core saccharide of the miharamycins would be derived from cyclization of a dihydroxyethyl or glycoaldehyde branch installed at C3. Our inspection of the *smh* cluster revealed two genes unique to this cluster, *smh26* and *smh27*, which encode an L-fuculose phosphate class II aldolase and a ribokinase, respectively. Considering that aldolases are capable of carbon-carbon bond formation and scission,[119,120] and phosphorylation can increase the nucleofugality of leaving groups,[121] these enzymes are good candidates for the cyclization the two-carbon branch installed by the *smh* transketolase, and three hypothetical pathways were envisioned (Figure 2.14).

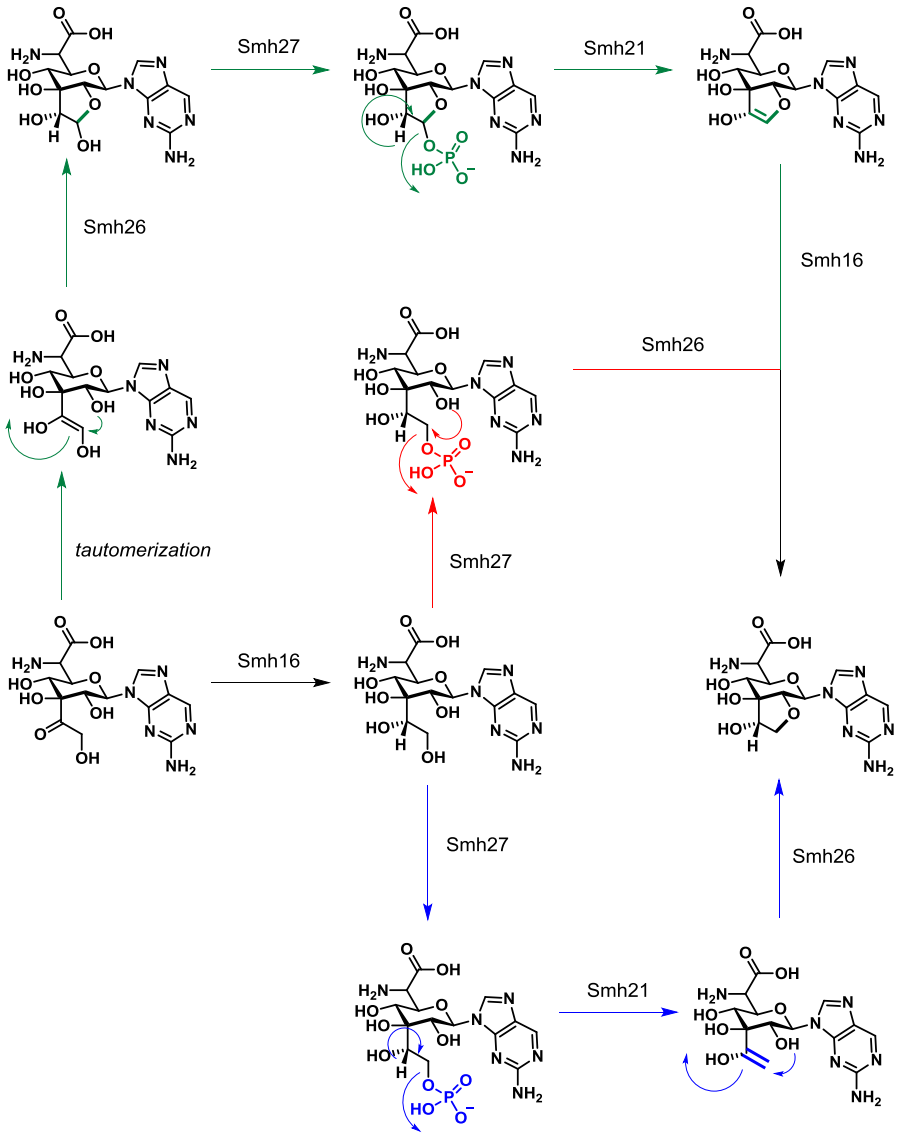



**Figure 2.14:** Hypothetical biosynthetic routes to the formation of the miharamycin A perhydrofuropyran.

In the first route, the terminal alcohol of C9 is phosphorylated by Smh27, facilitating Smh26-mediated nucleophilic attack on C9 by the C2-OH, resulting in expulsion of phosphate and formation of the furan ring (Figure 2.14, red route). In the second route, the C9 alcohol is also phosphorylated by Smh27, but then deprotonation at C8 eliminates phosphate and generates an electrophilic methylene which can be attacked

by the C2-OH in an Smh26-mediated process, likewise generating the furan ring (Figure 2.14, blue route). This pathway is complicated by the anticipated high $pK_a$ of the C8 protons, but in light of our proposal the methyltransferases Snv13 and Smh15 participate in the cyclization of the core triketide as unusual cyclases, this cyclization could require a third enzyme to fulfill a similar role. Supporting this possibility, the *smh* cluster encodes a second methyltransferase that is not shared with the *snv* cluster, Smh21, which could deprotonate C8 to yield the methylene intermediate, and Smh26 catalyzes the nucleophilic attack on this intermediate by the C2-OH. Finally, in the third hypothetical pathway, the glycoaldehyde branch could tautomerize to give an electrophilic alkene. Then, Smh26 mediates nucleophilic attack by the C2-OH on C8 (Figure 2.14, green route). Phosphorylation of the C9-OH by Smh27 could facilitate its elimination to yield a dihydrofuran, and then Smh16, the SDR, could reduce this intermediate to complete formation of the furan.

*2.3.2.2. Glycosylation/ Nucleobase Attachment*

If amipurimycin and the miharamycins are derived from a PKS, the nascent products of these enzymes would be lactones and, thus, incapable of glycosylation to yield nucleosides. Intriguingly, the *smh* and *snv* clusters contain the *myo*-inositol 2-dehydrogenases Smh4 and Snv1, respectively. *Myo*-inositol 2-dehydrogenase acts physiologically to oxidize *myo*-inositol, generating 2,4,6/3,5-pentahydroxycyclohexanone.[122] The similarity between the cyclohexanone product of *myo*-inositol 2-dehydrogenase and the proposed PKS precursors of amipurimycin and the miharamycins led us to believe Snv1 and Smh4 reduce the nascent triketide lactones to yield the hexopyranoside-containing core saccharides of these PNAs, which would then be capable of glycosylation.

As mentioned earlier, the *smh* and *snv* clusters both contain a three gene cassette composed of a CGA-synthase-like glycosyltransferase, an aminoglycoside resistance kinase, and a flavoprotein. While not always the case, the co-localization of these three genes in their respective clusters suggests they act in temporal proximity in the biosynthesis of amipurimycin and the miharamycins. This assertion is supported by the overlapping start and stop codons of these cassettes, which would be expected to lead to their transcription as polycistronic mRNAs.[123,124] Given the presence of the glycosyltransferase, we assumed these three genes are involved in the glycosylation of the core saccharides of the miharamycins and amipurimycin. The cytosyl-PNAs all rely on NDP-activated sugars for glycosylation,[9] but if the miharamycins and amipurimycin are PKS-derived, they would need to be activated. Because the cytosyl-PNAs usually rely on UDP-glucuronic acid, a physiological sugar, these pathways have no need for the inclusion of "activating" enzymes such as NDP-transferases. Supporting this logic, the biosynthetic gene cluster of amicetin, the one cytosyl-PNA not derived from UDP-glucuronic acid, includes a thymididyl transferase to activate its precursor, glucose-1-phosphate.[82]

The *snv* and *smh* clusters lack NDP-transferases, so one possibility is the core saccharides are activated by NDP-transferases recruited from elsewhere. However, the presence of the aminoglycoside resistance kinases Smh7 and Snv5 in close proximity to glycosyltransferases suggests these kinases play a role in the activation of a sugar substrate for glycosylation. The "true" glycosyltransferases typically require activation of their substrates by a high-energy substituted phosphate, e.g., pyrophosphate, acyl phosphate, triphosphate, but some types of glycosyltransferases are known to utilize monophosphorylated sugar substrates.[125] Thus, it is not unreasonable to suggest the kinases Snv5 and Smh7 phosphorylate the "anomeric" hydroxyls of the reduced amipurimycin and

miharamycins PKS-derived sugars to facilitate their attachment to 2-aminopurine or a precursor of this rare nucleobase.

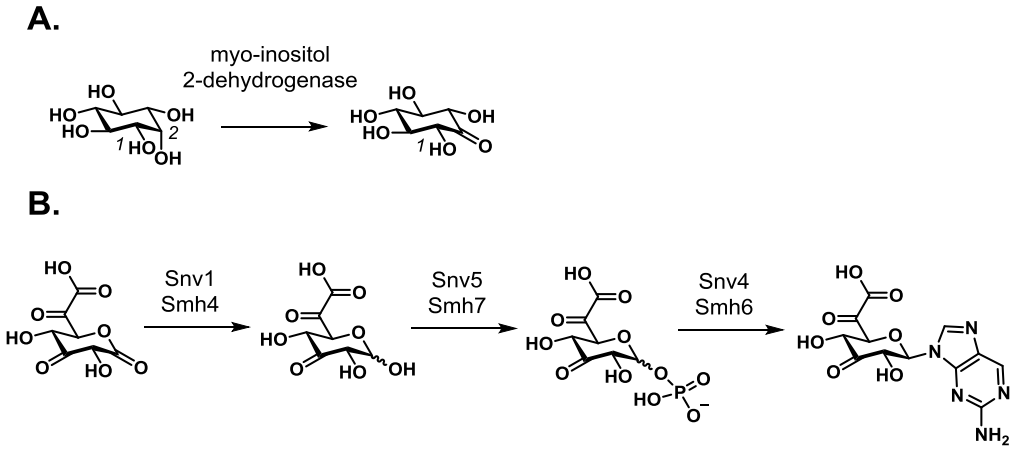



**Figure 2.15:** Myo-inositol 2-dehydrogenase reaction compared to reduction of a nascent PKS-derived pyran; putative glycosylation reaction sequence.

*2.3.2.3. 2-aminopurine Biosynthesis*

The biosynthetic origin of the 2-aminopurine base is not clear, but as mentioned earlier, the lack of a nucleoside hydrolase found in the biosynthetic gene clusters of other cytosyl-PNAs indicates the base could originate from a free purine precursor. Furthermore, neither the *smh* or *snv* clusters nor their genomes contain paralogous purine biosynthetic enzymes, a finding which does not support a parallel alternative de novo purine biosynthetic pathway. On the other hand, it is possible 2-aminopurine is derived from a physiological base. Purines lacking a C6 substituent are exceedingly rare in nature because the physiologically relevant purines are made from inosine, which has a C6 carbonyl.[126] One example of a naturally-occurring purine lacking a C6 substituent is nebularine, also known as purine ribonucleoside.[13] This natural product is produced by the fungus *Lepista nebularis* and the bacteria *Streptomyces yokosukanensis*[127] and *Microbiospora* sp.

SCC1779.[128] While the detailed mechanism has yet to be established, isotope-labeled precursor feeding studies [8-$^{14}$C]inosine, [8-$^{14}$C]adenosine, and [8-$^{3}$H]guanosine in *S. yokosukanensis* established adenosine as the precursor to nebularine, and that adenosine is incorporated without prior catabolism. The reaction of adenosine with partially purified enzyme from *S. yokosukanensis* was found to produce nebularine with concomitant production of hydroxylamine. Surprisingly, the partially purified protein seemed to contain no chromophoric cofactor, such as heme or a flavin, and the addition of exogenous cofactors or metal ions or EDTA had no effect on the production of nebularine.[127]

The flavoprotein found in the *smh* and *snv* clusters in the glycosylation cassette, Smh8 and Snv6, respectively, is a predicted flavin-dependent oxidoreductase of unknown function. The only common purine also featuring a C2-amino functional group is guanine, which makes it or one of its derivatives a reasonable hypothetical precursor to the unusual 2-aminopurine base found in the miharamycins and amipurimycin. As expected based on analyses using the NCBI Conserved Domain Database (CDD), Smh8 and Snv6 feature an extended Rossmann fold which is indicative of flavin or NAD(P)H binding.

Surprisingly, although not detected by BLASTP homology or CDD analyses, Smh8 and Snv6 also contain a derivative of the CXXCXXXC motif, which is indicative of iron-sulfur cluster binding.[129] If, indeed, 2-aminopurine is derived from guanine, the

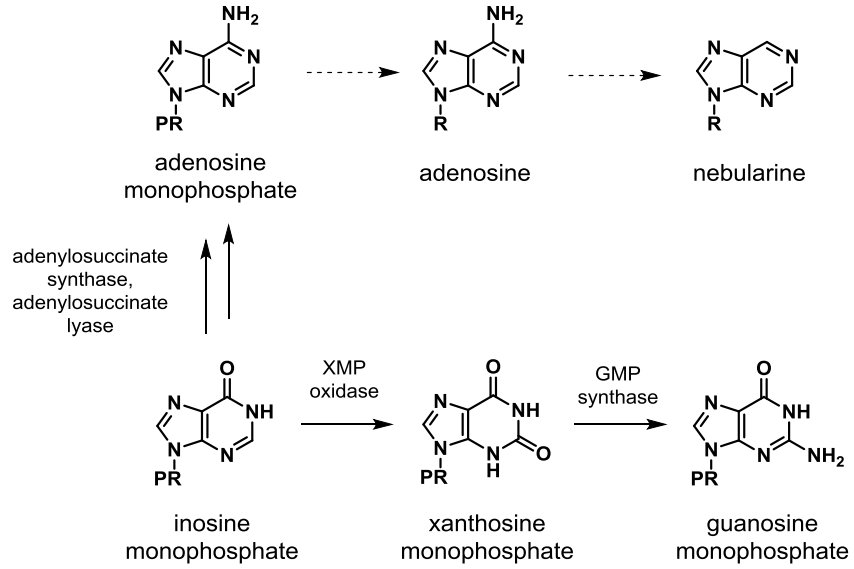



**Figure 2.16:** The biosynthesis of nebularine juxtaposed on *de novo* purine biosynthesis.

removal of the C6-carbonyl resembles the reverse reaction of xanthine oxidase.[126] However, these analyses alone provide no insight into the timing of the Smh8 or Snv6 reaction, and further investigation is needed to characterize the function of these oxidoreductases.

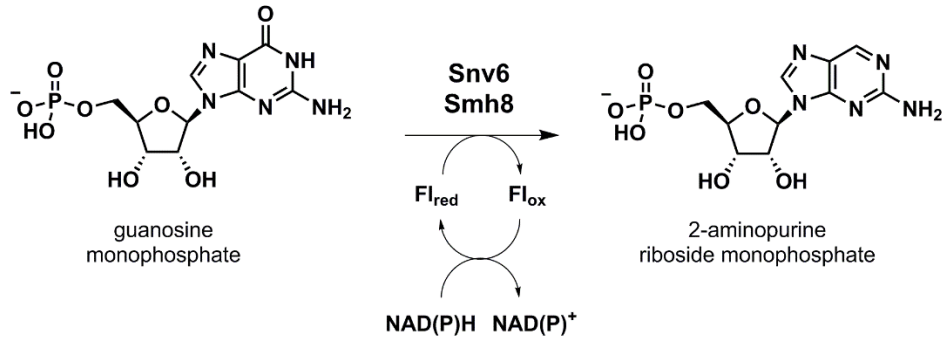



**Figure 2.17:** A hypothetical origin of 2-aminopurine from 5'-GMP.

*2.3.2.4. Amino Acids*

Similarly to other hPNAs, the smh and snv clusters encode ATP-grasp ligases (Smh 20 and Snv18) which can be reasonably assigned as the ligases attaching the pendant amino acids to their respective cores, as no other means of peptide bond formation can be found in the clusters.[10,109] Because the amino acids found in the structures of the miharamycin A and amipurimycin are non-proteinogenic, it was predicted that the *smh* and *snv* gene clusters would encode the enzymes necessary for their biosynthesis of their unique amino acids. This prediction seems to hold true for the cispentacin biosynthetic genes.

Flanking the core set of genes shared with the smh cluster, the snv cluster includes a set of genes, snvA-H, encoding a full set of close homologues of enzymes comprising the highly conserved type II PKS biosynthetic pathway of 2-carboxy-2-cyclopentanone (CPC), a precursor of coronafacic acid in the biosynthesis of coronatine.[74,130] Coronafacic acid is a virulence factor produced by several pathovars of Pseudomonas syringae, but the coronatine biosynthetic gene cluster from *Pseudomonas syringae glycinea* PG4180 is the most extensively studied.[74,76] Recently, a coronafacic acid-like biosynthetic gene cluster was also identified in *Streptomyces scabies*, the causative agent of common potato scab.[131] Coronafacic acid is composed of the unusual amino acid coronamic acid and coronafacic acid, a polyketide.[74,76,130] The dissociated PKS system composed of Cfa1-5 generates CPC, which serves as a starter unit for the multi-modular type I PKS system comprised of Cfa6 and Cfa7. While the biosynthesis of coronafacic acid has been investigated starting from CPC, very little work has been done regarding the biosynthesis of CPC itself.[76]

On the basis of labeled precursor feeding studies, 2-ketoglutarate semialdehyde is proposed to contribute four of the six carbon atoms of CPC, and malonate provides the remaining two carbon atoms. Interestingly, it is believed α-ketoglutarate semialdehyde originates

from primary metabolism or at least from an enzyme outside of the coronatine biosynthetic gene cluster, since no decarboxylase was identified in the cluster.[74]

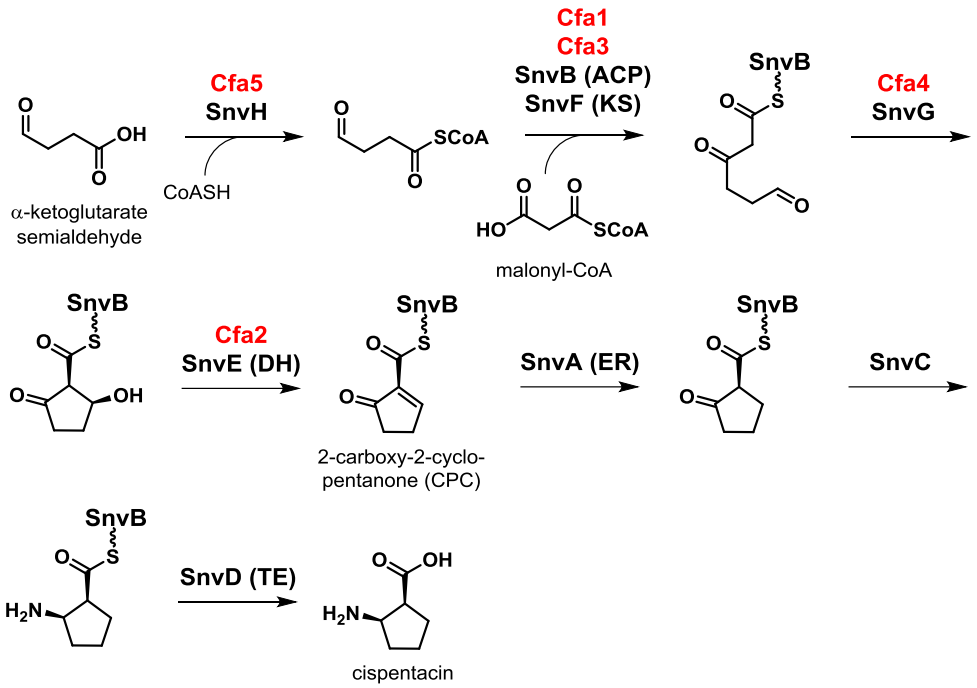



**Figure 2.18:** Proposed biosynthetic pathway with hypothetical route to cispentacin.

The biosynthesis of CPC begins with the conversion of 2-ketoglutarate semialdehyde into succinic semialdehyde CoA thioester by Cfa5, a fatty acid-CoA ligase, resulting in the loss of one equivalent of carbon dioxide. This activated substrate is then condensed with Cfa1(ACP)-linked malonate by the ketosynthase Cfa3. The resultant linear Cfa1-bound product is then proposed to undergo a Diekmann cyclization catalyzed by Cfa4. Cfa4 was originally assigned as the requisite cyclase even though its product has no homology to known cyclases because no other cyclase could be identified. Finally, the cyclized product is dehydrated by Cfa2, a dehydratase, yielding CPC.[74] This pathway is summarized in Figure 14, which includes the homologs found in the *snv* cluster. ACP-linked CPC is then

shuttled to Cfa6 for further elaboration into coronafacic acid. Interestingly, hypothetical reduction and transamination of CPC would yield cispentacin. In support of this hypothesis, the *snv* CPC sub-cluster was found to encode an enoyl reductase, SnvA, and an L-ornithine aminotransferase homolog, SnvC, enzymes not found in other identified CPC clusters. In addition, a unique thioesterase (SnvD) is included in the subcluster. Therefore, SnvA may represent the entry point to the cispentacin biosynthetic pathway, the reduced CPC can be transaminated by SnvC, and then cispentacin is released from its ACP by SnvD.

The biosynthesis of the $N^5$-hydroxyarginine found in the structure of miharamycin A is not as straightforward. Though the hydroxylated guanidine resembles a key intermediate in the reaction of nitric oxide synthase,[132] no nitric oxide synthases or similar enzymes could be clearly identified in the *smh* cluster. However, $N^5$-hydroxyarginine was also recently found as a component of argolaphos, a phosphonate natural product produced by *Streptomyces monomycini* NRRL B-24309.[3] Surprisingly, the reported argolaphos biosynthetic gene cluster encodes a homologue of Smh24, and these proteins are predicted YqcI/YcgG homoloues. Two lines of evidence suggest these enzymes may play a role in the biosynthesis of $N^5$-hydroxyarginine. First, our analysis using the RAST server suggests YqcI/YcgG is a conserved enzyme found in several species of *Bacillus*, and two different species of Bacillus have been reported to produce $N^5$-hydroxyarginine.[31,32] Secondly, our bioinformatic analysis of *Bacillus* strains in the RAST database shows YqcI/YcgG lies in a three gene cassette also containing an amino acid permease and an arginine deiminase shared between *Oceanobacillus iheyensis* HTE831 and *Bacillus* B-14905, *Bacillus vietnamensis*, and many other species of *Bacillus*. The *O. iheyensis* HTE831 cluster also includes a homolog of LysE, an efflux protein for basic amino acids, and *Bacillus cereus* VD131 and *Bacillus anthracis* have YqcI/YcgG homologs adjacent to a LysE homolog (*B.*

*anthracis* has two sets at different genomic loci). Although Smh24 lacks recognizable cofactor binding motifs, it may be an unusual monooxygenase. Furthermore, the weak homology between Agp12 and Smh24 (24.7% similarity, 17.7% identity) may reflect the structural dissimilarity between argolaphos and miharamycin A, suggesting the modification of arginine occurs after arginine is incorporated into these compounds.

*2.3.2.4. A Paradigm of Novel PNAs?*

The PKS machinery encoded in the *smh* and *snv* clusters is not found in any other known PNA gene clusters, but through the course of analysis, four additional highly similar putative biosynthetic gene clusters were identified in unannotated data in the NCBI database in the organisms *Streptomyces* sp. S-1868 (*ses* cluster), *Streptomyces varsoviensis* NRRL B-3589 (*svs* cluster), *Streptomyces canus* str. 299MFChir4.1 (*sca* cluster), and *Actinoplanes* sp. N902-109 (*apn* cluster). The products of these four "orphan" gene clusters are not known, but because they all share the same major subparts of the *snv* and *smh* clusters, we speculate these products are structurally similar.

The most surprising aspect of these clusters is their modularity and synteny, which suggests they are capable of producing compounds similar to amipurimycin and the miharmycins. Specifically, these clusters feature close homologues of the PKS and loading modules found in the *snv* and *smh* gene clusters, and three of the four gene clusters have an intact Fkb subcluster.

Interestingly, these clusters all have homologues of the *smh* and *snv* methyltransferases of unknown function, which indicates they play an important role in the biosynthesis of amipurimycin- and miharamycins-like compounds. The S.

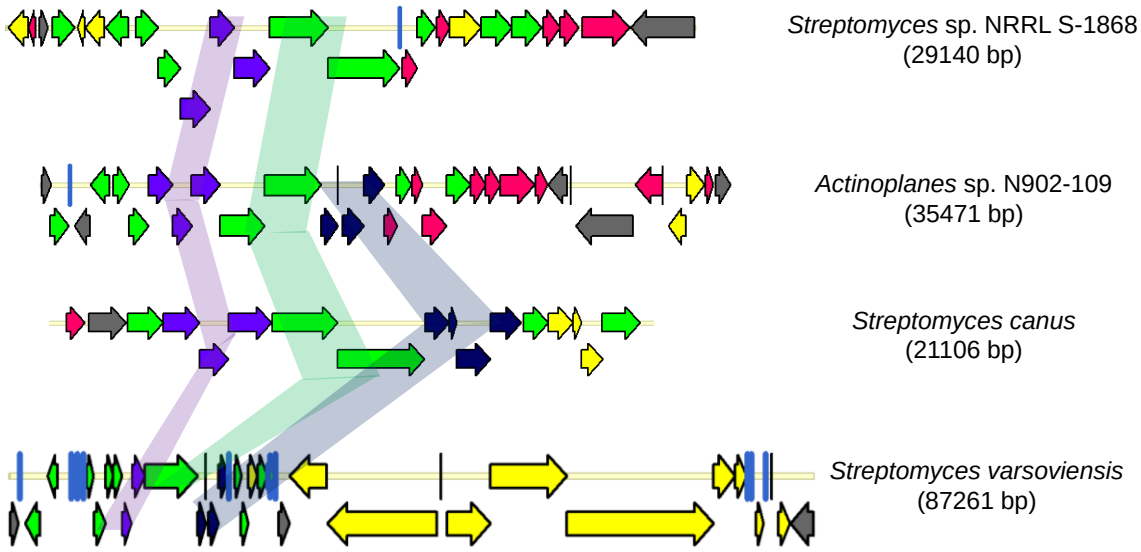



**Figure 2.19:** Organization of putative amipurimycin/ miharamycins-like PNA gene clusters

*varsoviensis* cluster appears to be the most divergent cluster. While it seems to have the core components shared by the other clusters, like the transketolase and Fkb subcluster, the monomodular PKS is fused with the putative loading module observed in the other clusters, and this locus also has a partial queosine cluster composed of QueC-E and GTP cyclohydrolase I homologs. Interestingly, the cluster lacks the branched-chain aminotransferase found in the other five clusters. Unlike the amino acid portion of amipurimycin, the probable product may have an α-hydroxy acid moiety instead of the amino acid-like portion of amipurimycin and the miharamycins. This is supported by the lack of an ATP-grasp ligase in the cluster. The divergent modular PKS subclusters seem to produce two different saturated fatty acids, which could be appended to the alpha-hydroxyl group. In all, the *S. varsoviensis* cluster may be capable of producing several products.

## 2.4. CONCLUSIONS

Sequencing of the *S. miharaensis* and *S. novoguineensis* genomes and comparative genomic analysis afforded identification of the putative miharamycins and amipurimycin biosynthetic gene clusters, *smh* and *snv*, respectively. The content of these clusters is remarkably conserved between the two producing organisms. Unexpectedly, the clusters contain enzymes typically associated with the biosynthesis of PKS-derived natural products. It is tantalizing to implicate this PKS machinery in the biosyntheses of the miharamycins and amipurimycin, but there is no precedent for such a pathway in the biosynthesis of carbohydrates. Additionally, the possibility these compounds are derived from enzymes typical of polyketide biosynthesis holds promise for the greater understanding of the versatility of PKS enzymology. Nevertheless, the data enabled the creation of a reasonable biosynthetic model, laying the groundwork for further investigation into the biosyntheses of these unusual PNAs. Finally, the *smh* and *snv* clusters may potentially be paradigmatic of a new class of PKS-derived carbohydrates, as four highly conserved gene clusters in four distinct bacterial species were newly identified. Several issues remain to be explored regarding the biosynthesis of amipurimycin and the miharamycins, but given the rarity of these compounds and their components, it is expected they will continue to fuel both future synthetic efforts and chemical biology studies.

## 2.5. REFERENCES

(1) Galperin, M. Y.; Koonin, E. V. From complete genome sequence to 'complete' understanding? *Trends Biotechnol.* **2010**, *28*, 398.

(2) Khosla, C. Quo vadis, enzymology? *Nat Chem Biol.* **2015**, *11*, 438.

(3) Ju, K. S.; Gao, J.; Doroghazi, J. R.; Wang, K. K.; Thibodeaux, C. J.; Li, S.; Metzger, E.; Fudala, J.; Su, J.; Zhang, J. K.; Lee, J.; Cioni, J. P.; Evans, B. S.;

Hirota, R.; Labeda, D. P.; van der Donk, W. A.; Metcalf, W. W. Discovery of phosphonic acid natural products by mining the genomes of 10,000 actinomycetes. *Proc Natl Acad Sci U S A.* **2015**, *112*, 12175.

(4) Katz, L.; Baltz, R. H. Natural product discovery: past, present, and future. *J Ind Microbiol Biotechnol.* **2016**.

(5) Richards, S. It's more than stamp collecting: how genome sequencing can unify biological research. *Trends Genet.* **2015**, *31*, 411.

(6) Kim, E.; Moore, B. S.; Yoon, Y. J. Reinvigorating natural product combinatorial biosynthesis with synthetic biology. *Nat Chem Biol.* **2015**, *11*, 649.

(7) Hertweck, C. The biosynthetic logic of polyketide diversity. *Angew Chem Int Ed Engl.* **2009**, *48*, 4688.

(8) Hertweck, C. Decoding and reprogramming complex polyketide assembly lines: prospects for synthetic biology. *Trends Biochem Sci.* **2015**, *40*, 189.

(9) Niu, G.; Tan, H. Nucleoside antibiotics: biosynthesis, regulation, and biotechnology. *Trends Microbiol.* **2015**, *23*, 110.

(10) Walsh, C. T.; Fischbach, M. A. Natural products version 2.0: connecting genes to molecules. *J Am Chem Soc.* **2010**, *132*, 2469.

(11) Walsh, C. T.; Zhang, W. Chemical logic and enzymatic machinery for biological assembly of peptidyl nucleoside antibiotics. *ACS Chem Biol.* **2011**, *6*, 1000.

(12) Rachakonda, S.; Cartee, L. Challenges in antimicrobial drug discovery and the potential of nucleoside antibiotics. *Curr Med Chem.* **2004**, *11*, 775.

(13) Isono, K. Nucleoside antibiotics: structure, biological activity, and biosynthesis. *J Antibiot (Tokyo).* **1988**, *41*, 1711.

(14) Isono, K. Current progress on nucleoside antibiotics. *Pharmacol Ther.* **1991**, *52*, 269.

(15) Rackham, E. J.; Gruschow, S.; Ragab, A. E.; Dickens, S.; Goss, R. J. Pacidamycin biosynthesis: identification and heterologous expression of the first uridyl peptide antibiotic gene cluster. *Chembiochem.* **2010**, *11*, 1700.

(16) Chen, W.; Huang, T.; He, X.; Meng, Q.; You, D.; Bai, L.; Li, J.; Wu, M.; Li, R.; Xie, Z.; Zhou, H.; Zhou, X.; Tan, H.; Deng, Z. Characterization of the polyoxin biosynthetic gene cluster from Streptomyces cacaoi and engineered production of polyoxin H. *J Biol Chem.* **2009**, *284*, 10627.

(17) Lauer, B.; Russwurm, R.; Schwarz, W.; Kalmanczhelyi, A.; Bruntner, C.; Rosemeier, A.; Bormann, C. Molecular characterization of co-transcribed genes from Streptomyces tendae Tu901 involved in the biosynthesis of the peptidyl

moiety and assembly of the peptidyl nucleoside antibiotic nikkomycin. *Mol Gen Genet.* **2001**, *264*, 662.

(18) Shubitz, L. F.; Trinh, H. T.; Perrill, R. H.; Thompson, C. M.; Hanan, N. J.; Galgiani, J. N.; Nix, D. E. Modeling nikkomycin Z dosing and pharmacology in murine pulmonary coccidioidomycosis preparatory to phase 2 clinical trials. *J Infect Dis.* **2014**, *209*, 1949.

(19) Stenland, C. J.; Lis, L. G.; Schendel, F. J.; Hahn, N. J.; Smart, M. A.; Miller, A. L.; von Keitz, M. G.; Gurvich, V. J. A practical and scalable manufacturing process for an anti-fungal agent, Nikkomycin Z. *Org Process Res Dev.* **2013**, *17*, 265.

(20) Shubitz, L. F.; Roy, M. E.; Nix, D. E.; Galgiani, J. N. Efficacy of Nikkomycin Z for respiratory coccidioidomycosis in naturally infected dogs. *Med Mycol.* **2013**, *51*, 747.

(21) Gunasekera, A.; Alvarez, F. J.; Douglas, L. M.; Wang, H. X.; Rosebrock, A. P.; Konopka, J. B. Identification of GIG1, a GlcNAc-induced gene in Candida albicans needed for normal sensitivity to the chitin synthase inhibitor nikkomycin Z. *Eukaryot Cell.* **2010**, *9*, 1476.

(22) Nix, D. E.; Swezey, R. R.; Hector, R.; Galgiani, J. N. Pharmacokinetics of nikkomycin Z after single rising oral doses. *Antimicrob Agents Chemother.* **2009**, *53*, 2517.

(23) Cone, M. C.; Petrich, A. K.; Gould, S. J.; Zabriskie, T. M. Cloning and heterologous expression of blasticidin S biosynthetic genes from Streptomyces griseochromogenes. *J Antibiot (Tokyo).* **1998**, *51*, 570.

(24) Tercero, J. A.; Espinosa, J. C.; Lacalle, R. A.; Jimenez, A. The biosynthetic pathway of the aminonucleoside antibiotic puromycin, as deduced from the molecular analysis of the pur cluster of Streptomyces alboniger. *J Biol Chem.* **1996**, *271*, 1579.

(25) Genilloud, O.; Gonzalez, I.; Salazar, O.; Martin, J.; Tormo, J. R.; Vicente, F. Current approaches to exploit actinomycetes as a source of novel natural products. *J Ind Microbiol Biotechnol.* **2011**, *38*, 375.

(26) Bachmann, B. O.; Van Lanen, S. G.; Baltz, R. H. Microbial genome mining for accelerated natural products discovery: is a renaissance in the making? *J Ind Microbiol Biotechnol.* **2014**, *41*, 175.

(27) Goto, T.; Toya, Y.; Kondo, T. Structure of amipurimycin, a new nucleoside antibiotic produced by Streptomyces novoguineensis. *Nucleic Acids Symp Ser.* **1980**, s73.

(28) Seto, H.; Koyama, M.; Ogino, H.; Tsuruoka, T.; Inouye, S.; Otake, N. The structures of novel nucleoside antibiotics, miharamycin A and miharamycin B. *Tetrahedron.* **1983**, *24*, 1805.

(29) Marcelo, F.; Jimenez-Barbero, J.; Marrot, J.; Rauter, A. P.; Sinay, P.; Bleriot, Y. Stereochemical assignment and first synthesis of the core of miharamycin antibiotics. *Chemistry.* **2008**, *14*, 10066.

(30) Fischer, B.; Keller-Schierlein, W.; Kneifel, H.; Konig, W. A.; Loeffler, W.; Muller, A.; Muntwyler, R.; Zahner, H. Metabolic products of microorganisms. 118. Delta-N-hydroxy-L-arginine, an amino acid antagonist from Nannizzia gypsea. *Archiv fur Mikrobiologie.* **1973**, *93*, 203.

(31) Maehr, H.; Blount, J. F.; Pruess, D. L.; Yarmchuk, L.; Kellett, M. Antimetabolite produced by microorganisms. VIII. N5-hydroxy-L-arginine, a new naturally occurring amino acid. *J. Antibiot.* **1973**, *26*, 284.

(32) Perlman, D.; Vlietinck, A. J.; Matthews, H. W.; Lo, F. F. Microbial production of vitamin B12 antimetabolites. I. N5-hydroxy-L-arginine from Bacillus cereus 439. *J. Antibiot.* **1974**, *27*, 826.

(33) Nemoto, A.; Hoshino, Y.; Yazawa, K.; Ando, A.; Mikami, Y.; Komaki, H.; Tanaka, Y.; Grafe, U. Asterobactin, a new siderophore group antibiotic from Nocardia asteroides. *J. Antibiot.* **2002**, *55*, 593.

(34) Makarieva, T. N.; Ogurtsova, E. K.; Denisenko, V. A.; Dmitrenok, P. S.; Tabakmakher, K. M.; Guzii, A. G.; Pislyagin, E. A.; Es'kov, A. A.; Kozhemyako, V. B.; Aminin, D. L.; Wang, Y. M.; Stonik, V. A. Urupocidin A: a new, inducing iNOS expression bicyclic guanidine alkaloid from the marine sponge Monanchora pulchra. *Org Lett.* **2014**, *16*, 4292.

(35) Dobashi, K.; Matsuda, N.; Hamada, M.; Naganawa, H.; Takita, T.; Takeuchi, T. Novel antifungal antibiotics octacosamicins A and B. I. Taxonomy, fermentation and isolation, physico-chemical properties and biological activities. *J Antibiot (Tokyo).* **1988**, *41*, 1525.

(36) Dobashi, K.; Naganawa, H.; Takahashi, Y.; Takita, T.; Takeuchi, T. Novel antifungal antibiotics octacosamicins A and B. II. The structure elucidation using various NMR spectroscopic methods. *J Antibiot (Tokyo).* **1988**, *41*, 1533.

(37) Konishi, M.; Nishio, M.; Saitoh, K.; Miyaki, T.; Oki, T.; Kawaguchi, H. Cispentacin, a new antifungal antibiotic. I. Production, isolation, physico-chemical properties and structure. *J Antibiot (Tokyo).* **1989**, *42*, 1749.

(38) Oki, T.; Hirano, M.; Tomatsu, K.; Numata, K.; Kamei, H. Cispentacin, a new antifungal antibiotic. II. In vitro and in vivo antifungal activities. *J Antibiot (Tokyo).* **1989**, *42*, 1756.

(39) Iwamoto, T.; Tsujii, E.; Ezaki, M.; Fujie, A.; Hashimoto, S.; Okuhara, M.; Kohsaka, M.; Imanaka, H.; Kawabata, K.; Inamoto, Y.; et al. FR109615, a new antifungal antibiotic from Streptomyces setonii. Taxonomy, fermentation, isolation, physico-chemical properties and biological activity. *J Antibiot (Tokyo).* **1990**, *43*, 1.

(40) Iwasa, T.; Kishi, T.; Matsuura, K.; Wakae, O. Streptomyces novoguineensis sp. Nov., an amipurimycin producer, and antimicrobial activity of amipurimycin. *J Antibiot (Tokyo).* **1977**, *30*, 1.

(41) Shomura, T.; Hamamoto, K.; Ohashi, T.; Amano, S.; Yoshida, J.; Moriyama, C.; Niida, T. New antibiotics, miharamycins A and B. II. Some biological characteristics of miharamycin. *Meiji Seika Kenkyu Nenpo.* **1967**, *9*, 5.

(42) Noguchi, T.; Yasuda, Y.; Niida, T.; Shomura, T. Inhibitory effects of Miharamycin A on the multiplication of plant viruses and the symptom development. *Annals of the Phytopathological Society of Japan.* **1968**, *34*, 323.

(43) Misra, A. The use of antibiotics for the control of plant virus diseases. *Journal of Plant Diseases and Protection.* **1977**, *84*, 244.

(44) Tsuruoka, T.; Yumoto, H.; Ezaki, N.; Niida, T. New antibiotics, miharamycins A and B. I. Isolation and characterization of miharamycins A and B. *Meiji Seika Kenkyu Nenpo.* **1967**, *9*, 1.

(45) Goodman, M. F.; Hopkins, R. L.; Lasken, R.; Mhaskar, D. N. The biochemical basis of 5-bromouracil- and 2-aminopurine-induced mutagenesis. *Basic Life Sci.* **1985**, *31*, 409.

(46) Hopkins, R. L.; Goodman, M. F. Ribonucleoside and deoxyribonucleoside triphosphate pools during 2-aminopurine mutagenesis in T4 mutator-, wild type-, and antimutator-infected Escherichia coli. *J Biol Chem.* **1985**, *260*, 6618.

(47) Ronen, A. 2-Aminopurine. *Mutat Res.* **1980**, *75*, 1.

(48) Sowers, L. C.; Boulard, Y.; Fazakerley, G. V. Multiple structures for the 2-aminopurine-cytosine mispair. *Biochemistry.* **2000**, *39*, 7613.

(49) Stauffer, C. S.; Datta, A. Synthetic studies on amipurimycin: total synthesis of a thymine nucleoside analogue. *J Org Chem.* **2008**, *73*, 4166.

(50) Cachatra, V.; Almeida, A.; Sardinha, J.; Lucas, S. D.; Gomes, A.; Vaz, P. D.; Florencio, M. H.; Nunes, R.; Vila-Vicosa, D.; Calhorda, M. J.; Rauter, A. P. Wittig Reaction: Domino Olefination and Stereoselectivity DFT Study. Synthesis of the Miharamycins' Bicyclic Sugar Moiety. *Org Lett.* **2015**, *17*, 5622.

(51) Gelfand, M. S.; Cleveland, K. O. Ceftolozane/Tazobactam Therapy of Respiratory Infections due to Multidrug-Resistant Pseudomonas aeruginosa. *Clin Infect Dis.* **2015**, *61*, 853.

(52) McCarthy, K. Pseudomonas aeruginosa: evolution of antimicrobial resistance and implications for therapy. *Semin Respir Crit Care Med.* **2015**, *36*, 44.

(53) Sambrook, J.; Russell, D. W. *Molecular cloning : a laboratory manual*; Third ed.; Cold Spring Harbor Laboratory: Cold Spring Harbor, NY, 2001.

(54) Kieser, T.; Bibb, M. J.; Buttner, M. J.; Chater, K. F.; Hopwood, D. A. *Practical streptomyces genetics*; The John Innes Foundation: Norwich, U.K., 2000.

(55) Niida, T.; Yumoto, H.; Tsuruoka, T.; Hamamoto, K.; Shomura, J.; Ohashi, T.; Antibiotics Obtained from Streptomyces Miharaensis; U.S. Patent 3,678,159, July 18, 1972.

(56) Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal.* **2011**, *17*, 10.

(57) Schiemer, J.; Rev. A ed.; Tufts University Core Facility: 2014; Vol. 2014.

(58) Data Processing of Nextera Mate Pair Read Data on Illumina Sequencing Platforms. Illumina: 2012.

(59) Gordon, A. FASTX-Toolkit. Computer program distributed by the author. **2011**.

(60) Cock, P. J. A.; Fields, C. J.; Goto, N.; Heuer, M. L.; Rice, P. M. The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants. *Nucleic Acids Res.* **2010**, *38*, 1767.

(61) Zerbino, D. R. Using the Velvet de novo assembler for short-read sequencing technologies. *Curr Protoc Bioinformatics.* **2010**, *Chapter 11*, Unit 11 5.

(62) Zerbino, D. R.; Birney, E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* **2008**, *18*, 821.

(63) Zerbino, D. R.; McEwen, G. K.; Margulies, E. H.; Birney, E. Pebble and rock band: heuristic resolution of repeats and scaffolding in the velvet short-read de novo assembler. *PLoS One.* **2009**, *4*, e8407.

(64) Aziz, R. K.; Bartels, D.; Best, A. A.; DeJongh, M.; Disz, T.; Edwards, R. A.; Formsma, K.; Gerdes, S.; Glass, E. M.; Kubal, M.; Meyer, F.; Olsen, G. J.; Olson, R.; Osterman, A. L.; Overbeek, R. A.; McNeil, L. K.; Paarmann, D.; Paczian, T.; Parrello, B.; Pusch, G. D.; Reich, C.; Stevens, R.; Vassieva, O.; Vonstein, V.; Wilke, A.; Zagnitko, O. The RAST Server: rapid annotations using subsystems technology. *BMC Genomics.* **2008**, *9*, 75.

(65) Overbeek, R.; Olson, R.; Pusch, G. D.; Olsen, G. J.; Davis, J. J.; Disz, T.; Edwards, R. A.; Gerdes, S.; Parrello, B.; Shukla, M.; Vonstein, V.; Wattam, A. R.; Xia, F.; Stevens, R. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res.* **2014**, *42*, D206.

(66) Blin, K.; Medema, M. H.; Kazempour, D.; Fischbach, M. A.; Breitling, R.; Takano, E.; Weber, T. antiSMASH 2.0--a versatile platform for genome mining of secondary metabolite producers. *Nucleic Acids Res.* **2013**, *41*, W204.

(67) Darling, A. C.; Mau, B.; Blattner, F. R.; Perna, N. T. Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res.* **2004**, *14*, 1394.

(68) Darling, A. E.; Treangen, T. J.; Messeguer, X.; Perna, N. T. Analyzing patterns of microbial evolution using the mauve genome alignment system. *Methods Mol Biol.* **2007**, *396*, 135.

(69) Darling, A. E.; Tritt, A.; Eisen, J. A.; Facciotti, M. T. Mauve assembly metrics. *Bioinformatics.* **2011**, *27*, 2756.

(70) Lee, H.; Tang, H. Next-generation sequencing technologies and fragment assembly algorithms. *Methods Mol Biol.* **2012**, *855*, 155.

(71) Comin, M.; Schimd, M. Assembly-free genome comparison based on next-generation sequencing reads and variable length patterns. *BMC Bioinformatics.* **2014**, *15 Suppl 9*, S1.

(72) Chin, F. Y.; Leung, H. C.; Yiu, S. M. Sequence assembly using next generation sequencing data--challenges and solutions. *Sci China Life Sci.* **2014**, *57*, 1140.

(73) Compeau, P. E.; Pevzner, P. A.; Tesler, G. How to apply de Bruijn graphs to genome assembly. *Nat Biotechnol.* **2011**, *29*, 987.

(74) Rangaswamy, V.; Jiralerspong, S.; Parry, R.; Bender, C. L. Biosynthesis of the Pseudomonas polyketide coronafacic acid requires monofunctional and multifunctional polyketide synthase proteins. *Proc Natl Acad Sci U S A.* **1998**, *95*, 15469.

(75) Rangaswamy, V.; Mitchell, R.; Ullrich, M.; Bender, C. Analysis of genes involved in biosynthesis of coronafacic acid, the polyketide component of the phytotoxin coronatine. *J Bacteriol.* **1998**, *180*, 3330.

(76) Strieter, E. R.; Koglin, A.; Aron, Z. D.; Walsh, C. T. Cascade reactions during coronafacic acid biosynthesis: elongation, cyclization, and functionalization during Cfa7-catalyzed condensation. *J Am Chem Soc.* **2009**, *131*, 2113.

(77) Bentley, S. D.; Chater, K. F.; Cerdeno-Tarraga, A. M.; Challis, G. L.; Thomson, N. R.; James, K. D.; Harris, D. E.; Quail, M. A.; Kieser, H.; Harper, D.; Bateman, A.; Brown, S.; Chandra, G.; Chen, C. W.; Collins, M.; Cronin, A.; Fraser, A.; Goble, A.; Hidalgo, J.; Hornsby, T.; Howarth, S.; Huang, C. H.; Kieser, T.; Larke, L.; Murphy, L.; Oliver, K.; O'Neil, S.; Rabbinowitsch, E.; Rajandream, M. A.; Rutherford, K.; Rutter, S.; Seeger, K.; Saunders, D.; Sharp, S.; Squares, R.; Squares, S.; Taylor, K.; Warren, T.; Wietzorrek, A.; Woodward, J.; Barrell, B. G.;

Parkhill, J.; Hopwood, D. A. Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2). *Nature.* **2002**, *417*, 141.

(78) Cone, M. C.; Yin, X.; Grochowski, L. L.; Parker, M. R.; Zabriskie, T. M. The blasticidin S biosynthesis gene cluster from Streptomyces griseochromogenes: sequence analysis, organization, and initial characterization. *Chembiochem.* **2003**, *4*, 821.

(79) Rackham, E. J.; Gruschow, S.; Goss, R. J. Revealing the first uridyl peptide antibiotic biosynthetic gene cluster and probing pacidamycin biosynthesis. *Bioeng Bugs.* **2011**, *2*, 218.

(80) Wu, J.; Li, L.; Deng, Z.; Zabriskie, T. M.; He, X. Analysis of the mildiomycin biosynthesis gene cluster in Streptoverticillum remofaciens ZJU5119 and characterization of MilC, a hydroxymethyl cytosyl-glucuronic acid synthase. *Chembiochem.* **2012**, *13*, 1613.

(81) Niu, G.; Li, L.; Wei, J.; Tan, H. Cloning, heterologous expression, and characterization of the gene cluster required for gougerotin biosynthesis. *Chem Biol.* **2013**, *20*, 34.

(82) Zhang, G.; Zhang, H.; Li, S.; Xiao, J.; Zhu, Y.; Niu, S.; Ju, J.; Zhang, C. Characterization of the amicetin biosynthesis gene cluster from Streptomyces vinaceusdrappus NRRL 2363 implicates two alternative strategies for amide bond formation. *Appl Environ Microbiol.* **2012**, *78*, 2393.

(83) Lin, C. I.; McCarty, R. M.; Liu, H. W. The biosynthesis of nitrogen-, sulfur-, and high-carbon chain-containing sugars. *Chem Soc Rev.* **2013**, *42*, 4377.

(84) Cheng, L.; Chen, W.; Zhai, L.; Xu, D.; Huang, T.; Lin, S.; Zhou, X.; Deng, Z. Identification of the gene cluster involved in muraymycin biosynthesis from Streptomyces sp. NRRL 30471. *Mol Biosyst.* **2011**, *7*, 920.

(85) Zeng, Y.; Kulkarni, A.; Yang, Z.; Patil, P. B.; Zhou, W.; Chi, X.; Van Lanen, S.; Chen, S. Biosynthesis of albomycin delta(2) provides a template for assembling siderophore and aminoacyl-tRNA synthetase inhibitor conjugates. *ACS Chem Biol.* **2012**, *7*, 1565.

(86) Kaysser, L.; Siebenberg, S.; Kammerer, B.; Gust, B. Analysis of the liposidomycin gene cluster leads to the identification of new caprazamycin derivatives. *Chembiochem.* **2010**, *11*, 191.

(87) Funabashi, M.; Baba, S.; Nonaka, K.; Hosobuchi, M.; Fujita, Y.; Shibata, T.; Van Lanen, S. G. The biosynthesis of liposidomycin-like A-90289 antibiotics featuring a new type of sulfotransferase. *Chembiochem.* **2010**, *11*, 184.

(88) Knapp, S.; Morriello, G. J.; Nandan, S. R.; Emge, T. J.; Doss, G. A.; Mosley, R. T.; Chen, L. Assignment of the liposidomycin diazepanone stereochemistry. *J Org Chem.* **2001**, *66*, 5822.

(89) Chan, Y. A.; Podevels, A. M.; Kevany, B. M.; Thomas, M. G. Biosynthesis of polyketide synthase extender units. *Nat Prod Rep.* **2009**, *26*, 90.

(90) Chan, Y. A.; Thomas, M. G. Formation and characterization of acyl carrier protein-linked polyketide synthase extender units. *Methods Enzymol.* **2009**, *459*, 143.

(91) Liou, G. F.; Khosla, C. Building-block selectivity of polyketide synthases. *Curr Opin Chem Biol.* **2003**, *7*, 279.

(92) Du, L.; Lou, L. PKS and NRPS release mechanisms. *Nat Prod Rep.* **2010**, *27*, 255.

(93) Bretschneider, T.; Heim, J. B.; Heine, D.; Winkler, R.; Busch, B.; Kusebauch, B.; Stehle, T.; Zocher, G.; Hertweck, C. Vinylogous chain branching catalysed by a dedicated polyketide synthase module. *Nature.* **2013**, *502*, 124.

(94) Heine, D.; Bretschneider, T.; Sundaram, S.; Hertweck, C. Enzymatic polyketide chain branching to give substituted lactone, lactam, and glutarimide heterocycles. *Angew Chem Int Ed Engl.* **2014**, *53*, 11645.

(95) Heine, D.; Sundaram, S.; Bretschneider, T.; Hertweck, C. Twofold polyketide branching by a stereoselective enzymatic Michael addition. *Chem Commun (Camb).* **2015**, *51*, 9872.

(96) Kusebauch, B.; Busch, B.; Scherlach, K.; Roth, M.; Hertweck, C. Polyketide-chain branching by an enzymatic Michael addition. *Angew Chem Int Ed Engl.* **2009**, *48*, 5001.

(97) Sugimoto, Y.; Ishida, K.; Traitcheva, N.; Busch, B.; Dahse, H. M.; Hertweck, C. Freedom and constraint in engineered noncolinear polyketide assembly lines. *Chem Biol.* **2015**, *22*, 229.

(98) Chen, H.; Du, L. Iterative polyketide biosynthesis by modular polyketide synthases in bacteria. *Appl Microbiol Biotechnol.* **2016**, *100*, 541.

(99) Shen, B. Polyketide biosynthesis beyond the type I, II and III polyketide synthase paradigms. *Curr Opin Chem Biol.* **2003**, *7*, 285.

(100) Chan, Y. A.; Boyne, M. T., 2nd; Podevels, A. M.; Klimowicz, A. K.; Handelsman, J.; Kelleher, N. L.; Thomas, M. G. Hydroxymalonyl-acyl carrier protein (ACP) and aminomalonyl-ACP are two additional type I polyketide synthase extender units. *Proc Natl Acad Sci U S A.* **2006**, *103*, 14349.

(101) Chan, Y. A.; Thomas, M. G. Recognition of (2S)-aminomalonyl-acyl carrier protein (ACP) and (2R)-hydroxymalonyl-ACP by acyltransferases in zwittermicin A biosynthesis. *Biochemistry.* **2010**, *49*, 3667.

(102) Chen, D.; Zhang, Q.; Cen, P.; Xu, Z.; Liu, W. Improvement of FK506 production in Streptomyces tsukubaensis by genetic enhancement of the supply of

unusual polyketide extender units via utilization of two distinct site-specific recombination systems. *Appl Environ Microbiol.* **2012**, *78*, 5093.

(103) Kevany, B. M.; Rasko, D. A.; Thomas, M. G. Characterization of the complete zwittermicin A biosynthesis gene cluster from Bacillus cereus. *Appl Environ Microbiol.* **2009**, *75*, 1144.

(104) Luo, Y.; Ruan, L. F.; Zhao, C. M.; Wang, C. X.; Peng, D. H.; Sun, M. Validation of the intact zwittermicin A biosynthetic gene cluster and discovery of a complementary resistance mechanism in Bacillus thuringiensis. *Antimicrob Agents Chemother.* **2011**, *55*, 4161.

(105) Thibodeaux, C. J.; Melancon, C. E.; Liu, H. W. Unusual sugar biosynthesis and natural product glycodiversification. *Nature.* **2007**, *446*, 1008.

(106) Thibodeaux, C. J.; Melancon, C. E., 3rd; Liu, H. W. Natural-product sugar biosynthesis and enzymatic glycodiversification. *Angew Chem Int Ed Engl.* **2008**, *47*, 9814.

(107) Fischer, K.; Llamas, A.; Tejada-Jimenez, M.; Schrader, N.; Kuper, J.; Ataya, F. S.; Galvan, A.; Mendel, R. R.; Fernandez, E.; Schwarz, G. Function and structure of the molybdenum cofactor carrier protein from Chlamydomonas reinhardtii. *J Biol Chem.* **2006**, *281*, 30186.

(108) Tokunaga, H.; Kojima, M.; Kuroha, T.; Ishida, T.; Sugimoto, K.; Kiba, T.; Sakakibara, H. Arabidopsis lonely guy (LOG) multiple mutants reveal a central role of the LOG-dependent pathway in cytokinin activation. *Plant J.* **2012**, *69*, 355.

(109) Kuroha, T.; Tokunaga, H.; Kojima, M.; Ueda, N.; Ishida, T.; Nagawa, S.; Fukuda, H.; Sugimoto, K.; Sakakibara, H. Functional analyses of LONELY GUY cytokinin-activating enzymes reveal the importance of the direct activation pathway in Arabidopsis. *Plant Cell.* **2009**, *21*, 3152.

(110) Cai, W.; Goswami, A.; Yang, Z.; Liu, X.; Green, K. D.; Barnard-Britson, S.; Baba, S.; Funabashi, M.; Nonaka, K.; Sunkara, M.; Morris, A. J.; Spork, A. P.; Ducho, C.; Garneau-Tsodikova, S.; Thorson, J. S.; Van Lanen, S. G. The Biosynthesis of Capuramycin-type Antibiotics: IDENTIFICATION OF THE A-102395 BIOSYNTHETIC GENE CLUSTER, MECHANISM OF SELF-RESISTANCE, AND FORMATION OF URIDINE-5'-CARBOXAMIDE. *J Biol Chem.* **2015**, *290*, 13710.

(111) Goswami, A.; Van Lanen, S. G. Enzymatic strategies and biocatalysts for amide bond formation: tricks of the trade outside of the ribosome. *Mol Biosyst.* **2015**, *11*, 338.

(112) Zhao, Q.; He, Q.; Ding, W.; Tang, M.; Kang, Q.; Yu, Y.; Deng, W.; Zhang, Q.; Fang, J.; Tang, G.; Liu, W. Characterization of the azinomycin B biosynthetic gene cluster revealing a different iterative type I polyketide synthase for naphthoate biosynthesis. *Chem Biol.* **2008**, *15*, 693.

(113) Peng, C.; Pu, J. Y.; Song, L. Q.; Jian, X. H.; Tang, M. C.; Tang, G. L. Hijacking a hydroxyethyl unit from a central metabolic ketose into a nonribosomal peptide assembly line. *Proc Natl Acad Sci U S A.* **2012**, *109*, 8540.

(114) Chen, H.; Guo, Z.; Liu, H.-w. Biosynthesis of yersiniose: attachment of the two-carbon branched-cain is catalyzed by a thiamine pyrophosphate-dependent flavoprotein. *J Am Chem Soc.* **1998**, *120*, 11796.

(115) Hanessian, S.; Huang, G.; Chenel, C.; Machaalani, R.; Loiseleur, O. Total synthesis of N-malayamycin A and related bicyclic purine and pyrimidine nucleosides. *J Org Chem.* **2005**, *70*, 6721.

(116) Hanessian, S.; Marcotte, S.; Machaalani, R.; Huang, G. Total synthesis and structural confirmation of malayamycin A: a novel bicyclic C-nucleoside from Streptomyces malaysiensis. *Org Lett.* **2003**, *5*, 4277.

(117) Li, W.; Csukai, M.; Corran, A.; Crowley, P.; Solomon, P. S.; Oliver, R. P. Malayamycin, a new streptomycete antifungal compound, specifically inhibits sporulation of Stagonospora nodorum (Berk) Castell and Germano, the cause of wheat glume blotch disease. *Pest Manag Sci.* **2008**, *64*, 1294.

(118) Isono, F.; Inukai, M. Mureidomycin A, a new inhibitor of bacterial peptidoglycan synthesis. *Antimicrob Agents Chemother.* **1991**, *35*, 234.

(119) Houge-Frydrych, C. S.; Readshaw, S. A.; Bell, D. J. SB-219383, a novel tyrosyl tRNA synthetase inhibitor from a Micromonospora sp. II. Structure determination. *J Antibiot (Tokyo).* **2000**, *53*, 351.

(120) Samland, A. K.; Wang, M.; Sprenger, G. A. MJ0400 from Methanocaldococcus jannaschii exhibits fructose-1,6-bisphosphate aldolase activity. *FEMS Microbiol Lett.* **2008**, *281*, 36.

(121) Schneider, S.; Sandalova, T.; Schneider, G.; Sprenger, G. A.; Samland, A. K. Replacement of a phenylalanine by a tyrosine in the active site confers fructose-6-phosphate aldolase activity to the transaldolase of Escherichia coli and human origin. *J Biol Chem.* **2008**, *283*, 30064.

(122) Stirling, C. J. M. Leaving groups and nucleofugality in elimination and other organic reactions. *Acc Chem Res.* **1979**, *12*, 198.

(123) Ramaley, R.; Fujita, Y.; Freese, E. Purification and properties of Bacillus subtilis inositol dehydrogenase. *J Biol Chem.* **1979**, *254*, 7684.

(124) Lawrence, J. G. Gene organization: selection, selfishness, and serendipity. *Annu Rev Microbiol.* **2003**, *57*, 419.

(125) Lawrence, J. Selfish operons: the evolutionary impact of gene clustering in prokaryotes and eukaryotes. *Curr Opin Genet Dev.* **1999**, *9*, 642.

(126) Coutinho, P. M.; Deleury, E.; Davies, G. J.; Henrissat, B. An evolving hierarchical family classification for glycosyltransferases. *J Mol Biol.* **2003**, *328*, 307.

(127) Zhang, Y.; Morar, M.; Ealick, S. E. Structural biology of the purine biosynthetic pathway. *Cell Mol Life Sci.* **2008**, *65*, 3699.

(128) Brown, E. G.; Konuk, M. Biosynthesis of nebularine (purine 9-B-D-ribofuranoside) involves enzymic release of hydroxylamine from adenosine. *Phytochemistry.* **1995**, *38*, 61.

(129) Cooper, R.; Horan, A. C.; Gunnarsson, I.; Patel, M.; Truumees, I. Nebularine from a novel Microbispora sp. *J Ind Microbiol* **1986**, *1*, 275.

(130) Layer, G.; Gaddam, S. A.; Ayala-Castro, C. N.; Ollagnier-de Choudens, S.; Lascoux, D.; Fontecave, M.; Outten, F. W. SufE transfers sulfur from SufS to SufB for iron-sulfur cluster assembly. *J Biol Chem.* **2007**, *282*, 13342.

(131) Bender, C. L.; Liyanage, H.; Palmer, D.; Ullrich, M.; Young, S.; Mitchell, R. Characterization of the genes controlling the biosynthesis of the polyketide phytotoxin coronatine including conjugation between coronafacic and coronamic acid. *Gene.* **1993**, *133*, 31.

(132) Bignell, D. R.; Seipke, R. F.; Huguet-Tapia, J. C.; Chambers, A. H.; Parry, R. J.; Loria, R. Streptomyces scabies 87-22 contains a coronafacic acid-like biosynthetic cluster that contributes to plant-microbe interactions. *Mol Plant Microbe Interact.* **2010**, *23*, 161.

(133) Griffith, O. W.; Stuehr, D. J. Nitric oxide synthases: properties and catalytic mechanism. Annu Rev Physiol. 1995, 57, 707.

# Chapter 3: Initial Characterization of Amipurimycin and Miharamycins Biosynthesis

## 3.1. INTRODUCTION

In the last chapter, the putative miharamycins and amipurimycin biosynthetic gene clusters, *smh* and *snv*, from the genomes of *Streptomyces miharaensis* and *Streptomyces novoguineensis*, respectively, were identified through a bioinformatics approach. As is typical for the majority of natural product biosynthesis studies lately,[1-3] the obtainment of the *smh* and *snv* gene clusters preceded the collection of any significant biochemical data regarding the biosyntheses of the miharamycins and ampurimycins. To guide our investigations into the functions of the *smh* and *snv* clusters, we turned to the literature regarding the biosynthesis of peptidyl nucleoside antibiotics (PNAs).

While the miharamycins and amipurimycin are PNAs, they are best considered as part of a smaller subset of PNAs, the hexopyranosyl PNAs (hPNAs). The hPNAs are built upon a pyran-containing core saccharide resembling a hexose, and, based on the classification scheme presented in Chapter 1, the known hPNAs generally belong to the "intact" group of PNAs whose core saccharides derived from an intact sugar other than ribose which is not lengthened by the addition of any more carbon atoms. The hPNAs whose biosynthetic gene clusters had been sequenced include the cytosyl nucleosides blasticidin S,[4] mildiomycin,[5] arginomycin,[6] gougerotin,[7] and amicetin.[8] As discussed earlier, these compounds result from the condensation UDP-glucuronic acid and cytosine; with the exception of amicetin, which results from the condensation of a deoxysugar, TDP-amicetose, and cytosine. Thus, besides a glycosyltransferase to mediate this condensation, the gene clusters also include a nucleoside hydrolase to provide free cytosine.[8,9]

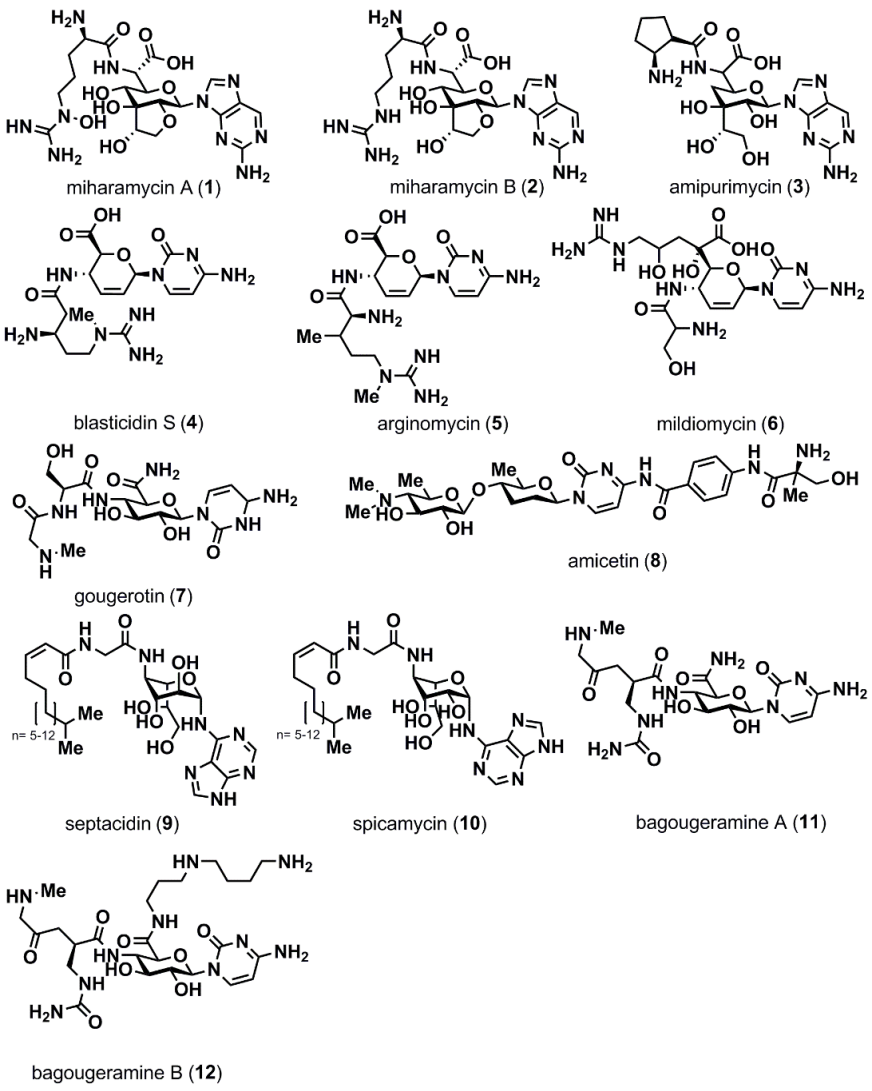



**Figure 3.1:** Structures of hexopyranosyl peptidyl nucleoside antibiotics.

Superficially, the miharamycins and amipurimycin seem to be members of hPNA, likely utilizing an intact-type biosynthetic route. However, as presented in Chapter 2, the content of the *smh* and *snv* clusters belies these designations. While the absence of the hydrolase is not too surprising considering the miharamycins and amipurimycin are not

cytosyl nucleosides, the lack of homologs of the glycosyltransferases in their putative gene clusters was unexpected.

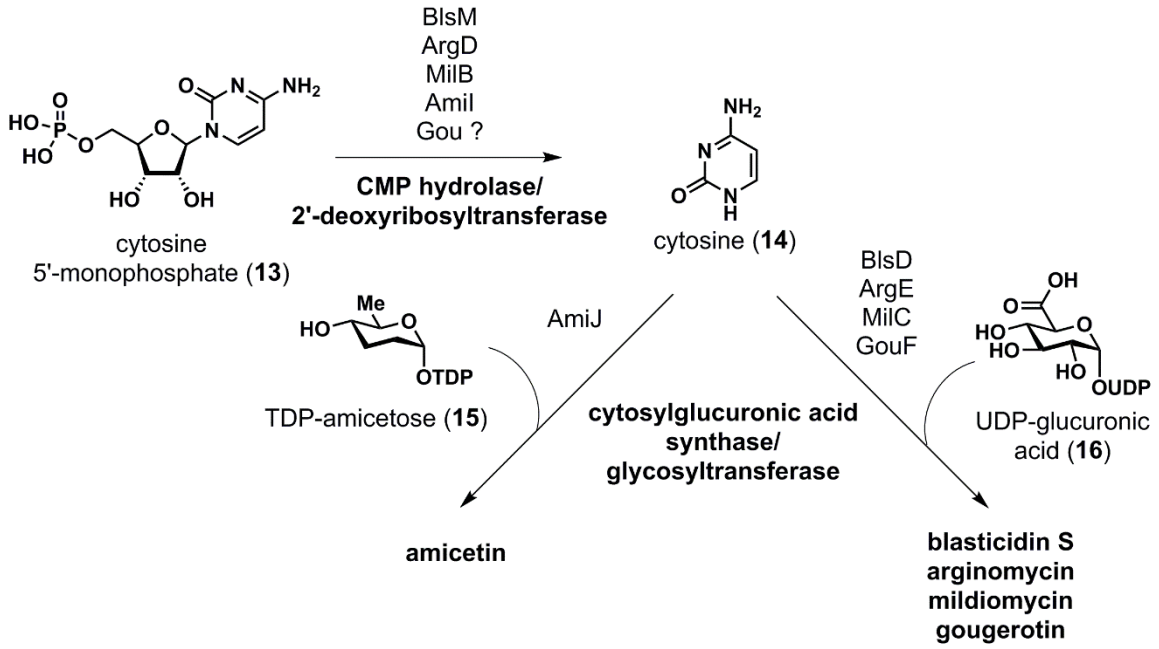



**Figure 3.2:** Overview of hexopyranosyl nucleoside formation pathways

Furthermore, unlike any of the hPNA gene clusters—or of any PNA gene clusters identified to date[7,10]—the *smh* and *snv* clusters are dominated by a type I polyketide synthease (PKS) as well as subclusters required for the biosynthesis of hydroxymalonate-based PKS extender units. These findings, along with the absence of enzymes involved in deoxy sugar biosynthesis, cast doubt on the involvement of these clusters in the biosynthesis of the miharamycins and amipurimycin. In this chapter, protocols for the production and isolation of miharamycins and amipurimycin as well as evidence for the involvement of the snv cluster in the production of amipurimycin are presented, and the results of initial attempts to observe the function of Snv4 and Snv6 are detailed.

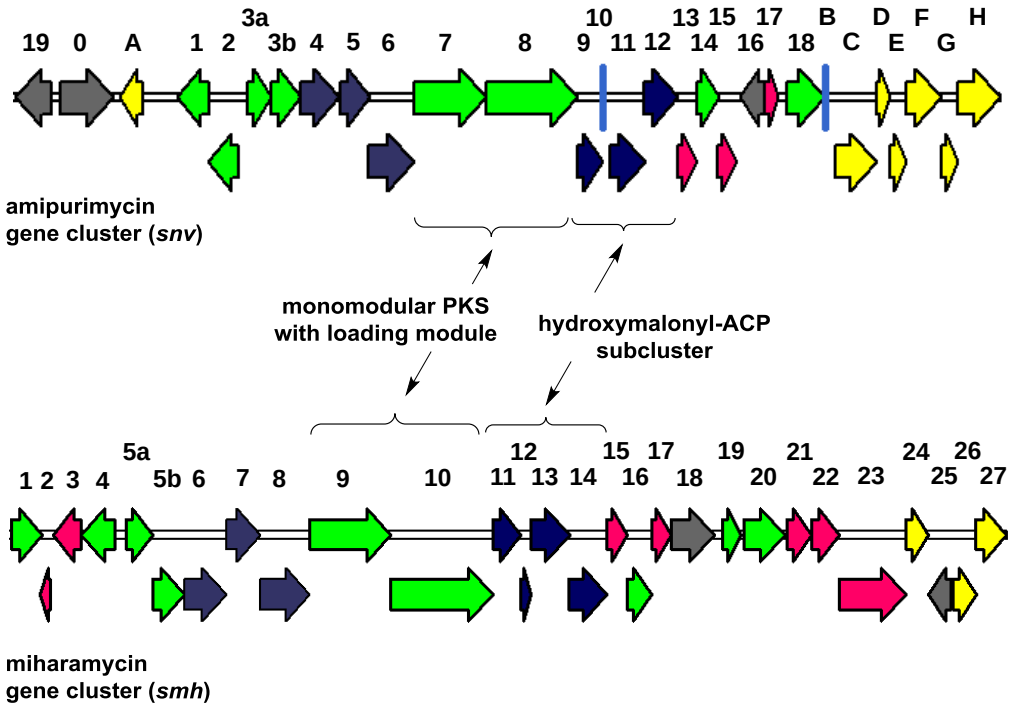



**Figure 3.3:** Organization of the snv and smh gene clusters.

## 3.2. MATERIALS AND METHODS

### *3.2.1. General*

All chemicals and reagents were purchased from Sigma-Aldrich (St. Louis, MO) or Fisher Scientific (Pittsburg, PA) and used without further purification unless otherwise noted. Laboratory disposables such as microcentrifuge tubes, petri dishes, and conical tubes were obtained from Fisher Scientific or VWR (Radnor, PA). With exceptions as noted, E. coli cloning and culture techniques as well as techniques relating to the manipulation of recombinant DNA were performed according to standard protocols.[11] Similarly, all Streptomyces manipulation methods and fermentations were performed according to standard protocols unless otherwise noted.[12] Restriction endonucleases (REN), T4 DNA ligase kit, and calf intestinal alkaline phosphatase (CIAP) were obtained from New England Biolabs (NEB) (Ipswich, MA), *E. coli* DH5α was obtained from

Bethesda Research Laboratories (Bethesda, MD), *E. coli* S17-1 ATCC 47055 was acquired from American Type and Culture Collection (Manassas, VA), and *E. coli* XL-1 Blue MRF' and *E. coli* BL21-Gold(DE3) were purchased from Stratagene (now Agilent, Santa Clara, CA). Sequencing vector pCRBlunt and 1 KB Plus DNA Ladder were products of Invitrogen (Carlsbad, CA). pET28b(+) vector DNA was obtained from Novagen (Madison, WI). EconoSpin® Mini Spin columns and buffers from Epoch Life Science (Sugar Land, TX) were used for mini-scale DNA isolation, DNA gel extraction, and PCR/ REN digestion reaction clean-up procedures according to the QIAgen *Miniprep Handbook, Fourth Edition*.[13] DNeasy Blood and Tissue kits and Ni-NTA resin were from QIAgen (Valencia, CA). All reagents and apparatus for SDS-PAGE were obtained from Bio-Rad (Hercules, CA).

*3.2.2. Growth and Production of Compounds and Feeding*

*Streptomyces miharaensis* Niida ATCC 19440 was obtained from American Type and Tissue Culture (Manassas, VA), and *Streptomyces novoguineensis* Iwasa CBS 199.78 was obtained from the CBS-KNAW Fungal Biodiversity Centre (Utrecht, The Netherlands). Both strains were maintained as described in Chapter 2. For the production of amipurimycin and miharamycins, 10 mL cultures of sterile seed medium (pH 7.0),[14] composed of 2% glucose, 1% Bacto® peptone from BD (Franklin Lakes, NJ) or peptone from Fisher Scientific, 0.5% meat extract from BD, 0.5% NaCl, and deionized (DI) H2O, in capped test tubes (25 x 150 mm) containing glass beads were inoculated with 0.2 mL mycelial suspension of either S. miharaensis or S. novoguineensis, and these cultures were grown at 28 °C with rotary shaking at 220 RPM in an Innova® 4200 Incubator-Shaker from New Brunswick Scientific (Edison, NJ) or a MaxQ® 6000 Incubator-Shaker from Thermo Scientific (Waltham, MA) for 2-3 days. Then, 1 mL of seed culture was used to inoculate

50 mL of sterile production medium,[15] composed of 4% sucrose, 3% soybean meal from Callahan's General Store (Austin, TX) or Choripdong® soybean flour (plain or roasted type, as available) from Seoul Trading, Inc. (Englewood, NJ), 2% Bacto® malt extract from BD or AMRESCO® malt extract from VWR, 0.6% NaCl, and DI H2O, sterilized by autoclave in a 1 L baffled flask covered with two milk filter discs from KenAG (Ashland, OH). This culture was grown for 4-5 days at 28 °C, with rotary shaking at 220 RPM, in the same incubator used for the seed cultures. Culture conditions were designed to approximate the effects of sparged oxygen.[16,17]

For culture feeding studies, 2-day-old production cultures were aseptically supplemented at 2 mM or other concentrations as needed using 0.2 μm-filter sterilized (cellulose acetate syringe filter from VWR) concentrated additive stock solutions and allowed to grow as before for an additional 2-3 days. The stocks were made fresh immediately before use. Additives were obtained from Sigma-Aldrich and included: L-arginine, L-ornithine, L-lysine, L-norleucine, 2-aminopurine, guanine, guanosine, 5'-guanosine monophosphate, 2-deoxy 5'-guanosine monophosphate, adenine, adenosine, thiamine, *O*-(carboxymethyl)-hydroxylamine, 6-diazo-5-oxo-L-norleucine, 1-aminobenzotriazole, and oxythiamine.

*3.2.3. Isolation of Compounds*

The protocol for isolation of miharamycins was based on published protocols for its isolation[14,15] and incorporated some elements from protocol for the isolation of nebularine.[18] Production cultures of *S. miharaensis* were harvested after 4-5 days of growth by transferring the culture medium into a sterile 50 mL conical tube from VWR and centrifugation at 5,000 x *g* in an Avanti® J-E centrifuge and JS-5.3 swinging bucket rotor from Beckman Coulter (Brea, CA) for 20 min at 4 °C. The supernatant was transferred to

a new sterile 50 mL conical tube containing 6 g pre-swollen Dowex® ($H^+$ form) MAC3 resin from Sigma-Aldrich, and this tube was spun in a Tube Rotator 60448 from Scientific Equipment Products (Aston, PA) at 4 °C overnight. Then, the supernatant was poured off, and the resin was rinsed with copious DI $H_2O$. The resin was transferred to a 500 mL Erlenmeyer flask containing a magnetic stir bar, and 100 mL 0.5 N aqueous HCl were added. The resin was stirred in HCl on a magnetic stir plate overnight at 4 °C, and then the supernatant was collected and adjusted to pH 3.0-4.0 using NaOH. The pH-adjusted extract was then stirred with 0.5 g DARCO® -100 mesh activated charcoal powder from Sigma-Aldrich at 4 °C overnight. The charcoal was collected by centrifugation at 5,000 x *g* in a sterile 50 mL conical tube and rinsed three times with DI $H_2O$. The charcoal pellet was then extracted with 50% aqueous acetone by stirring in a 500 mL Erlenmeyer flask at 4 °C overnight. The charcoal slurry was then filtered through a pad of Celite® 545 from Sigma-Aldrich in a sintered glass fritted funnel under vacuum. The filtrate was concentrated to dryness by rotary evaporation under reduced pressure in an R-114 Rotary Evaporator from Büchi Labortechnik AG (Flawil, Switzerland). The resultant residue was dissolved in <10 mL DI $H_2O$ and lyophilized in a sterile 50 mL conical tube. The crude powder was dissolved in 1 mL DI $H_2O$ for immediate use or stored at -80 °C until later use.

A "mini" scale isolation procedure was also developed to isolate miharamycins. To do so, 0.5 mL production culture supernatant was vortexed with 60 mg moist Dowex® MAC3 resin ($H^+$ form) in a 1.65 mL microcentrifuge tube from VWR, and then the supernatant was discarded. The resin was vortexed twice with 1 mL DI $H_2O$, and then washed by vortexing with 1 mL 0.5 N HCl. The acidic eluate was transferred to a new microcentrifuge tube containing 10 mg activated charcoal. After vortexing, the supernatant was removed by centrifugation in a 5415R tabletop microcentrifuge from Eppendorf

(Hamburg, Germany). The charcoal pellet was washed twice with 1 mL DI $H_2O$ by vortexing and centrifugation, and then the pellet was eluted with 1 mL 50% aqueous acetone by vortexing. The supernatant was collected in a new microcentrifuge tube and evaporated to dryness at ~55 °C using a SpeedVac® SC-100 from Savant (now Thermo Scientific). The residues were then stored at -80 °C until use.

Alternatively, miharamycins could be collected though open column chromatography with Dowex® MAC3 resin (6 g per 50 mL original culture volume) using isocratic elution with 0.5 N aqueous HCl. Miharamycin-containing fractions were detected with the Sakaguchi reaction which was performed as published before,[19,20] except 30% NaOH was used to obviate the need for neutralizing column fractions. Briefly, a 10 µL sample was first mixed with 2 µL ice cold 30% NaOH, then 2 µL ice cold 0.02% ethanolic α-naphthol were added, and the sample was mixed well by trituration. Color development was accomplished by the addition of 1 µL ice cold 0.2% sodium hypobromite in 5% NaOH. The appearance of an intense red or pink color indicated a positive reaction. The pooled fractions were then neutralized with NaOH and lyophilized in a 50 mL conical tube. The resultant residue was dissolved in 1 mL DI $H_2O$ and further purified using an Oasis MCX 1 mL SPE cartridge from Waters (Milford, MA) as follows. After conditioning the column with 1 mL each of MeOH and DI $H_2O$ by gravity flow, 1 mL sample composed of 0.5 mL lyophilate and 0.5 mL 3.4% $H_3PO_4$ was applied to the column. Then, the column was rinsed with 1 mL each of 0.1 N HCl and MeOH, and it was eluted with 5% $NH_4OH$ in MeOH.

For the isolation of amipurimycin, the culture supernatant was collected as for the miharamycins and transferred to a new 50 mL conical tube containing 0.5 g activated charcoal. This tube was spun in a Tube Rotator 60448 from Scientific Equipment Company (Aston, PA) overnight at 4 °C. Then, the charcoal was collected, rinsed, and

eluted as for the miharamycins. The elution was also concentrated, lyophilized, and stored as for the miharamycins.

*3.2.4. HPLC Analysis of Culture Extracts*

Extracts from cultures of *S. miharaensis* or *S. novoguineensis* were analyzed by HPLC using a Beckman Coulter System Gold® 125 Solvent Module and 166 UV Detector Module with 32 Karat HPLC Workstation Software from Beckman Coulter at ambient temperature and 314 nm or 223 nm. The samples were chromatographed on a Microsorb® MV 100-5 (250 x 4.6 mm) $C_{18}$ column from Varian (now Agilent) or ChromSep cartridge system with a Microsorb® MV 100-5 (250 x 4.6 mm) $C_{18}$ cartridge and ChromSep $C_{18}$ guard column (10 x 3 mm) by Varian from Agilent (Santa Clara, CA) as well as a Zorbax ODS 5 µm semi-preparative column (9.4 x 250 mm) from Agilent. Prior to analysis, samples were filtered through a 0.2 µm syringe filter. The binary mobile phase was composed of 0.1% (v/v) trifluoroacetic acid (TFA) from Chem-Impex International, Inc. (Wood Dale, IL) in DI H2O (solvent A) and 0.095% (v/v) TFA in acetonitrile (solvent B). Solvents were filtered through a 0.2 µm cellulose acetate filter (solvent A) or a PTFE 0.2 µm filter (solvent B) from EMDMillipore (Billerica, MA) prior to the addition of TFA. The elution program began with 0% solvent B and was carried out with a flow rate of 1 mL/min for analytical scale or 3 mL/min for semi-preparative scale as follows: at 5 minutes solvent B was increased to 2.5% over 2.5 min, at 15 min solvent B was increased to 5% over 2.5 min, then at 25 minute solvent B was increased to 100% over 3 min to flush the column, and then at 30 min solvent B was decreased to 0% over 3 minutes, and the column was re-equilibrated at 0% solvent B for 15 min. Detection of the miharamycins was first accomplished by comparison with an authentic standard of 70:30 miharamycin A:B mixture, a kind gift provided by Dr. Makoto Ohyama and Dr. Masato Tani of Meiji Seika

Kaisha, Ltd. (Tokyo, Japan). Mass spectrometry was carried out at the Mass Spectrometry and Proteomics Facility of the Department of Chemistry and Biochemistry at the University of Texas at Austin, Austin, TX.

*3.2.5. Bioassay to Detect Production of Miharamycins*

A disc-diffusion bioassay to detect production of the miharamycins was developed based on the published protocols.[14,15,21] A loopful of *Pseudomonas aeruginosa* PA01 glycerol stock, a kind gift from the laboratory of Dr. Marvin Whiteley of the University of Texas at Austin, was streaked on LB agar and grown overnight at 37 °C. The next day, a sterile toothpick was used to pick a single colony of *P. aeruginosa* and inoculate a 10 mL culture of LB broth in a sterile 50 mL conical tube. This culture was grown overnight in a shaking water bath at 220 RPM and 37 °C. The next day, this culture was used to inoculate the sterile, molten assay medium, composed of 0.5% dextrose, 0.5% peptone, 1.5% agar, and DI $H_2O$ (pH 7.0 prior to sterilization), at a 1:1000 ratio. After mixing well, 20-25 mL plates were made. After drying, sterile 9 mm Whatman® antibiotic assay discs from Sigma-Aldrich were aseptically applied to the surface of the plates using flamed forceps. Culture extracts were then applied to the discs, and after drying, the plates were incubated at 37 °C overnight. Zones of inhibition were measured as the width between the edge of the disc and the edge of the cleared zone on one side of the disc.

*3.2.6. Quantification of Miharamycins Production Following Putative Precursor Feeding*

Estimation of miharamycins concentration in production medium were made using a 2-aminopurine standard curve. A stock solution of 5 mg/mL 2-aminopurine nitrate salt from Sigma-Aldrich was made in DI $H_2O$, and serial dilutions were made in DI $H_2O$. Aliquots of serial dilutions were analyzed by HPLC using the same elution methods and

apparatus as used for analysis of culture extracts, and peak area integration was performed using 32 Karat HPLC Workstation Software.

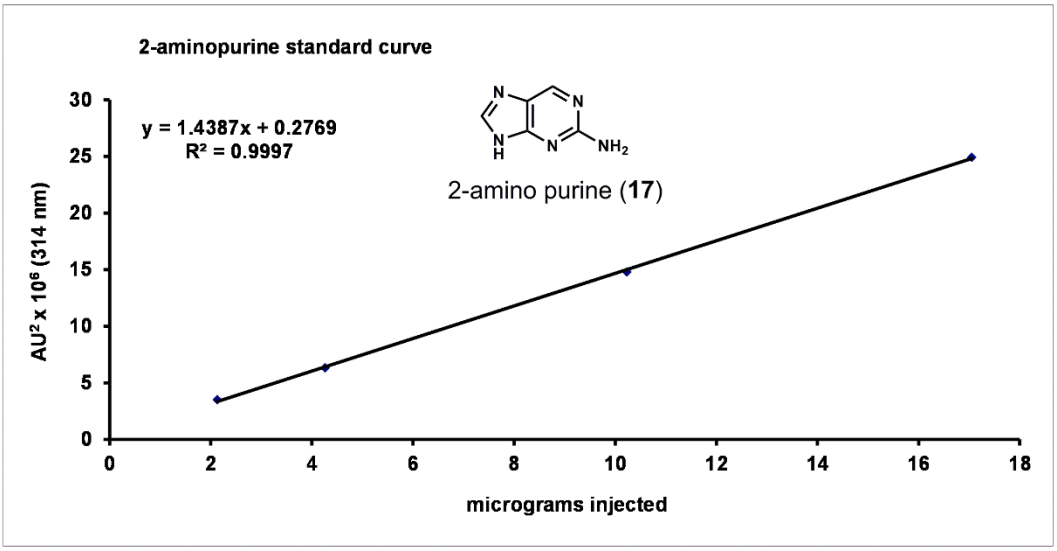



**Figure 3.4:** 2-amino purine standard curve

Production cultures of *S. miharaensis* were prepared and grown as described and fed after two days of growth with 5'-guanosine monophosphate (3 mM), L-ornithine (3 mM), or 5'-guanosine monophosphate and L-ornithine (3 mM each). Culture samples of 0.5 mL were aseptically collected at roughly 24 h time intervals for 150 h, and these samples were extracted using the mini-scale protocol described earlier. The residues were dissolved in 0.5 mL DI $H_2O$ and analyzed by HPLC as described earlier. The peak areas corresponding to miharamycin A and miharamycin B were integrated as described for the 2-aminopurine standard curve.

*3.2.7. Interruption of Snv8*

The internal 999 base-pair (bp) fragment of *snv8* ("snv8Int") was amplified using genomic DNA isolated from *S. novoguineensis* as described in Chapter 2 with the

following oligonucleotide primers obtained from Sigma-Aldrich: snv8IntFwd: 5'-TCTAGAGGGCGGGCTCGGACACCTG-3' and snv8IntRev: 5'-AAGCTTGCGAGCACACCG AGCATG -3'. These oligonucleotides have engineered *XbaI* and *HindIII* restriction endonuclease (REN) sites (underlined sequences), respectively. Polymerase chain reaction (PCR) was carried out with KOD HotStart® Polymerase from EMDMillipore (Billerica, MA) according to the manufacturer's instructions and using thin-walled PCR tubes. A Thermocycler Gradient® from Eppendorf was used and activation of the polymerase was accomplished with a 2 min incubation at 95 °C, and then 31 cycles composed of sequential incubations at 95 °C for 20 sec, the lowest melting temperature of the primer pair for 10 sec, and then 70 °C for 10 sec/kbp length of DNA target. A final annealing was carried out at 70 °C for 10 min. The PCR product was purified by agarose gel electrophoresis, and the gel-purified DNA was cloned into pCRBlunt vector using *E. coli* DH5α with T4 DNA ligase. The recombinant plasmids were sequenced at the Core Facilities of the Institute for Cell and Molecular Biology at the University of Texas at Austin. Correctly sequenced plasmid DNA was digested with XbaI and HindIII, and the desired DNA product was isolated through agarose gel electrophoresis. The fragment was ligated into pKC1139,[22] a kind gift from Dr. Leonard Katz (formerly of KOSAN Biosciences, Hayward, CA), digested with the same REN to yield plasmid pKCsnv8Int. Plasmid pKCsnv8Int was then transformed into chemically competent *E. coli* S17-123 for conjugal transfer to *S. novoguineensis* following a method derived from published protocols.[12,22,24]

To prepare *S. novoguineensis* for conjugation, 1 mL of 2-3 day old culture in seed medium was spread on a mannitol soya plate, and this plate was incubated at 30 °C for 1 week to grow a well-sporulated, confluent lawn. A single colony of *E. coli* S17-1 containing pKCsnv8Int was used to make a 5 mL culture in LB broth supplemented with

50 μg/mL apramycin in a sterile 50 mL conical tube, and this culture was grown at 37 °C overnight. *S. novoguineensis* spores were scraped from a confluent sporulating lawn using a flamed inoculating loop and deposited in 1 mL sterile LB in a sterile 1.65 mL microcentrifuge tube. This spore suspension was thoroughly vortexed and then incubated at 50 °C for 10-15 min. An aliquot of 1 mL of *E. coli* S17-1 culture containing plasmid pKCsnv8Int was centrifuged in a tabletop centrifuge at 10,000 x g for 1 min, and after removing the supernatant, 1 mL sterile LB was added to the pellet, and the pellet was gently resuspended. This suspension was rinsed once more and then resuspended in 1 mL sterile LB. An aliquot of 50 μL of the rinsed, resuspended E. coli culture were added to 1 mL of the *S. novoguineensis* spore solution, and this mixture was thoroughly vortexed. For each replicate conjugation experiment, 200 μL of the mixture were spread on a dehydrated mannitol soya plate supplemented with 10 mM MgCl2, and the plate was stored at room temperature overnight. Each plate was overlaid with 2 mL sterile soft R2 agar[25] containing 2.5 mg apramycin and 1.25 mg nalidixic acid, and after the overlay solidified, the plates were incubated at room temperature until sporulating colonies appeared (~1 week). Each colony was transferred to a fresh mannitol soya plate containing 100 μg/mL apramycin and 50 μg/mL nalidixic acid, and the plates were incubated at 40 °C. Two more spore-to-spore transfers were performed using mannitol soya plates containing 100 μg/mL apramycin and with incubation at 40 °C.

To verify interruption of *snv8* and assay ability of the exconjugants to produce amipurimycin, a loopful of spores of an exconjugant strain were streaked on a mannitol soya plate containing 100 μg/mL apramycin, and this plate was incubate at 40 °C for 2-3 days. Then, culture tubes containing 10 mL seed medium as described earlier supplemented with 100 μg/mL apramycin were inoculated with a single sporulating exconjugant colony. After growth as before at 30 °C, 1 mL of each culture was used for

isolation of genomic DNA using the DNeasy Blood and Tissue kit from QIAgen as described in Chapter 2, and 1 mL of each culture was used to inoculate 50 mL production medium as described earlier supplemented with 2 mM thiamine as described earlier and 100 μg/mL apramycin. These cultures were grown and extracted as before. The remaining seed culture volume was used to prepare frozen glycerol stocks as described in Chapter 2.

PCR analysis of the exconjugant genomic DNA was performed as described earlier using the oligonucleotides as primers, snv8Fd: 5'-TCTAGAATGTACGAGTTCGACCACG-3' and snv8Rv: 5'-AAGCTTTCATGACGCTGGGCTC-3' according to the KOD HostStart® polymerase manual with an annealing temperature of 60 °C and an extension time of 60 sec.

*3.2.8. Cosmid Library Construction*

For the isolation of genomic DNA for S. novoguineensis cosmid library preparation, a 0.3 mL portion of 2-3 day old culture grown in seed medium as described above was used to inoculate 100 mL of sterile seed medium supplemented with 0.5 % glycine. This culture was grown in a 1 L baffled flask at 30 °C with rotary shaking at 220 RPM for an additional 2-3 days. The culture was harvested in sterile 50 mL conical tubes by centrifugation at 5,000 x g for 20 min as described earlier. After removing the supernatants, the cell pellets were aseptically resuspended in sterile 10.3% sucrose and centrifuged again. Once more, the supernatants were removed, and the pellets were resuspended in 10.3% sucrose. After removing the supernatant, the pellets were stored at -80 °C until use. Genomic DNA was isolated by phenol:chloroform:isoamyl alcohol extraction and isopropanol precipitation according to standard protocols.[11,12]

*S. novoguineensis* genomic DNA isolated as described above was digested with Sau3AI under the following conditions: 50 μL DNA, 50 μL NEB buffer 4, 5 μL 100x NEB

BSA solution, and 387 µL water. The reaction was initiated by the addition of 8 µL Sau3AI, and after gently vortexing, the reaction was incubated at 37 °C. Aliquots of 100 µL were taken every 60 sec and quenched by the addition of 20 µL 6x DNA loading buffer. These samples were electrophoresed on a 0.8% agarose gel to gauge the reaction progress and extent of digestion; 2-3 min reaction time was deemed optimal. The digestion reaction was repeated and quenched at 2-3 min using 0.5 mL phenol:chloroform:isoamyl alcohol (25:24:1) saturated with 10 mM Tris, pH 8.0 and 1 mM EDTA from Sigma-Aldrich. After vortexing, the quenched reaction was centrifuged at 5,000 x *g* for 10 min, and then the upper/aqueous layer was transferred to a new microcentrifuge tube. An aliquot of 50 µL 3 M sodium acetate, pH 5.2, was added to the collected aqueous layer, and then 550 µL isopropyl alcohol were added, and the tube was thoroughly vortexed. After centrifugation at 16,100 x *g* for 45 min at 4 °C, the supernatant was removed, 500 µL 70% aqueous ethanol were added to the tube, and then it was centrifuged at 16,100 x g for 5 min. After removing the supernatant, the tube was allowed to air-dry for 1 h at room temperature, and then dried for 1 min using a Savant SpeedVac. Finally, the DNA pellet was dissolved in 15 µL elution buffer ("EB," 10 mM Tris-HCl, pH 8.5).

Cosmid vector pOJ446,[22] a kind gift from Dr. Leonard Katz (formerly of KOSAN Biosciences, Hayward, CA), was prepared for ligation by digestion with HpaI overnight at 37 °C according to the manufacturer's directions using ~20 µg DNA. The reaction was then treated with calf intestinal alkaline phosphatase (CIAP) from New England Biolabs according to the manufacturer's directions, and then purified using EconoSpin® Mini Spin columns and buffers from Epoch Life Science using the QIAgen PCR clean-up protocol,13 with elution using 50 µL EB. An aliquot of 20 µL of the spin column eluate (~6 µg DNA) from the last step was digested overnight with BamHI from New England Biolabs according to the manufacturer's directions; the reaction was purified using the QIAgen

PCR clean up protocol using EconoSpin® Mini Spin columns and buffers from Epoch Life Science with elution using 35 µL EB.

Ligation of the isolated, digested genomic DNA into the digested pOJ446 vector was carried out using T4 ligase as per the manufacturer's directions, except the reaction was incubated at 4 °C overnight. The next day, 1 µL 3 M sodium acetate, pH 5.5, was added to the reaction, followed by 11 µL isopropyl alcohol. The reaction was vortexed and centrifuged at 16,100 x *g* at 4 °C for 45 min, and then the supernatant was removed. An aliquot of 20 µL 70% aqueous ethanol was added to the microcentrifuge tube, and after centrifugation at 16,100 x *g* at 4 °C for 5 min, the supernatant was removed. The tube was allowed to air-dry at room temperature for 1 h, followed by drying in a Savant SpeedVac® for 1 min. The resultant pellet was dissolved in 12 µL EB.

The purified S. novoguineensis-pOJ446 ligation was packaged using the Giga Pack XL III kit from Stratagene (now Agilent) following the manufacturer's directions, and then the prepared packaging extract was stored at 4 °C. The recipient *E. coli* cells for each transfection were freshly prepared as follows. A loopful of *E. coli* XL-1 Blue MRF' glycerol stock was freshly streaked on an LB agar plate containing 12.5 µg/mL tetracycline, and the plate was incubated overnight at 37 °C. The next day, a singly colony was used to inoculate 5 mL LB liquid medium containing 12.5 µg/mL tetracycline, and this culture was allowed to grow overnight in a shaking water bath at 37 °C. The next day, the overnight culture was used to inoculate 25 mL fresh LB liquid made 0.2 % (v/v) maltose and 10 mM $MgSO_4$ at a 1:100 dilution, without tetracycline, and this new culture was grown at 37 °C in a shaking water bath until $OD_{600}$= 0.5-0.8. The cells were then treated as per the manufacturer's directions and used for transfection as follows. To assay the optimal transfection conditions, an initial trial was performed for each new packaging extract. Various amounts of the packaging extract were diluted to 130 µL using SM

buffer,[11] and 200 µL prepared *E. coli* XL-1 Blue MRF' cells were added to each dilution. The reactions were gently inverted and then incubated at room temperature for 30 min. A total of 1 mL LB was then added to each reaction, and the reactions were incubated at 37 °C for 45 min. The reactions were quickly pelleted by centrifugation for 10 sec, and most of the supernatant was poured off. The remaining ~200 µL supernatant was used to resuspend the cell pellets, each suspension was spread on one pre-warmed LB agar plate containing 50 µg/mL apramycin, and the plates were incubated at 37 °C overnight.

The next day, each plate from the transfection was flooded with 2 mL LB, and the plates were shaken by hand to wash off all of the colonies. Approximately 1.5 mL bacterial suspension was recoverable from each plate; 0.5 mL was used for miniprep isolation of DNA, and 125 µL 80% glycerol was added to the remaining 1 mL to make a frozen glycerol stock. Seven cosmid pools were obtained.

*3.2.9. Isolation of Amipurimycin Cluster Cosmid*

Based on analysis of the putative amipurimycin gene cluster from S. novoguineensis identified in Chapter 2, peg4680 and peg4707 were tentatively identified as the 5'- and 3'-cluster boundaries, respectively. Oligonucleotide primer pairs corresponding to 325 bp and 336 bp fragments, respectively, of these open reading frames (ORFs) were obtained from Sigma-Aldrich with the following sequences: 4680f: 5'-ATGCGGGTCTTCGGAACGATGACCGCC-3'; 4680r: 5'-AGCAGGGACTCGAACCGGTAGGCGTC-3'; 4707f: 5'-CCACCGGAGACCTCTTCACCATCGAC-3'; and 4707r: 5'-AGGAACGCGGTCACTTGTACCAGTCGATG-3'. PCR was carried out as described earlier with an annealing temperature and extension time of 67 °C and 3 sec, respectively, for peg4680 and an annealing temperature and extension time of 64 °C and 3 sec,

respectively, for peg4707. The seven cosmid pools collected earlier were screened with both probes using PCR, and one pool displaying a strong band for both probes, "NP2," was selected for further screening. An aliquot of 1 μL of NP2 DNA was transformed into chemically competent *E. coli* XL-1 Blue MRF' cells, and the transformation mixture was plated on LB agar containing 50 μg/mL apramycin and incubated overnight at 37 °C. A total of 150 individual colonies from this transformation were transferred to a gridded "master plate" and used to make three sub-pools as follows. A sterile toothpick was used to pick up a single colony, the toothpick was streaked on the gridded plate, and then the toothpick was dipped in 5 mL sterile LB containing 50 μg/mL apramycin and then discarded. Every 50 colonies were dipped in the same tube of LB to create three 5 mL cultures representing 150 colonies. After growth overnight in a shaking water bath at 37 °C, the sub-pool cultures were harvested and the DNA was collected by miniprep with elution using 50 μL EB. These sub-pools were PCR screened with the peg4680 and peg4707 probes, and subpool B ("NP2B"), was selected for further screening. The NP2B master plate was used to create five pools of 10 cosmids each using the approach described earlier. NP2B subpool five, corresponding to cosmids 41 through 50, showed strong bands for both probes, so each of these cosmids were individually screened. To do so, ten PCR reactions were set up as before but lacking template DNA, then a sterile toothpick was used to pick up a single colony, and this toothpick was dipped in the respective PCR reaction; this approach was used to supply the template DNA for the remaining reactions. Finally, after screening for both probes, only cosmid 47 ("NP2B47") showed a band for each probe. To verify the identity of the observed bands from the NP2B47 cosmid, these bands were excised, purified, and cloned into pCRBlunt as described earlier for sequencing. The resulting cosmid was named pSNVNP2B47, and since it was positive for both probes, it was assumed to contain the whole putative amipurimycin biosynthetic gene cluster.

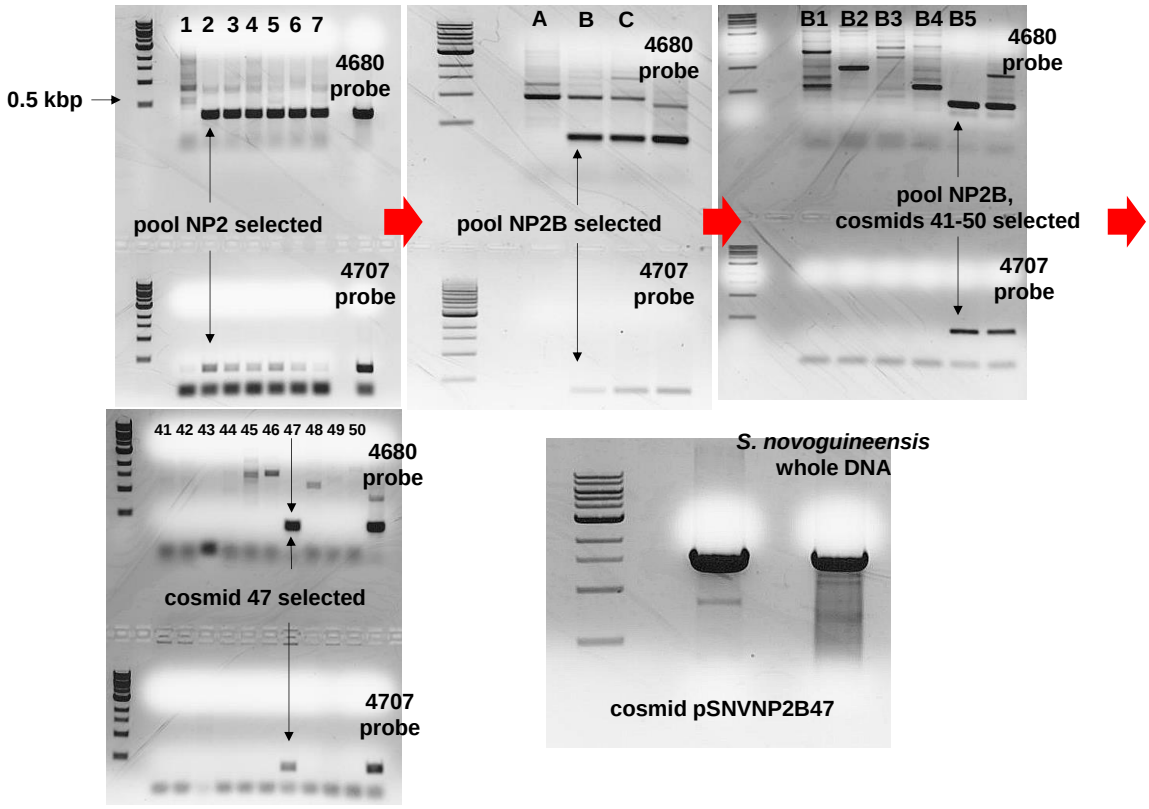



**Figure 3.5:** Iterative screening approach to isolation of cosmid pSNVNP2B47

*3.2.10. Cloning and Expression of Snv4 and Snv6*

The genes *snv4* and *snv6* were PCR amplified from cosmid pSNVNP2B47 as described earlier with the following oligonucleotide primer pairs with engineered 5'-*NdeI* and 3'-*NheI* sites (underlined) obtained from Sigma-Aldrich: snv4f: 5'-CATATGAGGATCTTCTCCAGTGCCCAG-3' and snv4r: 5'-GCTAGCTCACGACGTCCTCCCGGTGGGAAC-3'; and snv6f: 5'-CATATGTCTGACACCCCTGCCCGCAC-3' and snv6r: 5'-GCTAGCTCACAGACCCCCACCCGTCACGGTC-3'. The amplified genes were isolated by gel electrophoresis and cloned into pCRBlunt sequencing vector as described

earlier and sequenced. Correctly sequenced plasmids were digested with NdeI and NheI and cloned into pET28b(+) digested with the same REN to afford transcription of these genes as N-terminally His6-tagged constructs.[26] The recombinant plasmids containing *snv4* or *snv6* were transformed into chemically competent *E. coli* BL21-Gold(DE3) and plated on LB agar containing 50 μg/mL kanamycin. For transformants harboring either gene, a single colony was picked with a sterile toothpick and used to inoculate 20 mL sterile LB broth containing 50 μg/mL kanamycin in a sterile 50 mL conical tube. These cultures were grown overnight in a shaking water bath at 200 RPM and 37 °C overnight. The next day, these cultures were used to inoculate 4 L each of 1 L sterile LB in a 3.2 L Fernbach flask at a 1:250 dilution for protein production. These cultures were grown at 37 °C in an Innova® 4330 Refrigerated Incubator-Shaker from New Brunswick Scientific at 220 RPM until OD600 ~0.6-1.0. At that time, *snv4* cultures were induced with 0.05 mM isopropyl β-D-1-thiogalactopyranoside (IPTG), and *snv6* cultures were induced with 0.2 mM IPTG. At the time of induction, the incubator temperature was lowered to 24 °C, and growth was continued for 16-20 h. To harvest the cell pastes (typically ~5g/L culture), cultures were centrifuged at 4,500 x *g* at 4 °C for 20 min in an Avanti® J-E centrifuge with a JLA 10.5 fixed-angle rotor. The pastes were transferred to sterile 50 mL conical tubes and stored at -80 °C until use.

*3.2.11. Purification of Snv4 and Snv6*

To purify recombinant Snv4 and Snv6, the cell pastes were first thawed in room temperature water. The lysis, wash, and elution buffers contained imidazole as suggested[27] and were stored at 4 °C. The buffers for Snv4 purification were composed of 50 mM HEPES (pH 8.0), 300 mM NaCl, and 10% glycerol, and the buffers for Snv6 purification were composed of 50 mM Tris HCl (pH 8.0), 300 mM NaCl, 0.5 mM 1,4-dithiothreitol

(DTT), and 10% glycerol. Thawed cell pastes were resuspended and transferred to pre-chilled 250 mL stainless steel Griffin-style beakers from Polar Ware (Kiel, WI) with two cell paste volumes of the respective lysis buffer. Then, 1 mg of chicken egg-white lysozyme was added per mL of combined lysis and cell paste volume, and the mixture was incubated on ice for 1 h with intermittent stirring. The cell paste suspensions were then sonicated on ice using a Sonicator F550 from Fisher Scientific equipped with a large horn using 30 sec bursts at power setting "10" with 1 min cooling in between bursts. The suspensions were sonicated until no longer viscous (typically ~4-5 min total of sonication), and then they were centrifuged at 20,000 x *g* at 4 °C for 40 min in an Avanti® J-E centrifuge with a pre-chilled JS 25.50 fixed-angle rotor. The supernatants were then applied to pre-chilled, prepacked columns containing ~5.5 mL Ni-NTA resin pre-equilibrated with the respective lysis buffer and allowed to gravity flow. The columns were then washed with ~100 mL each of the respective wash buffer and eluted with ~15 mL each of the respective elution buffer. After checking fraction purity by SDS-PAGE, pooled fractions were loaded into washed Spectra/Por 2 (12-14 kDa molecular weight cutoff) dialysis tubing from Spectrum Laboratories (Rancho Dominguez, CA) and dialyzed against 3 L of dialysis buffer composed of the respective lysis buffers without imidazole at 4 °C. The dialyzed protein was concentrated to 5-7 mL using Amicon Ultra-15 Centrifugal Filter Units from EMDMillipore at 4 °C according to the manufacturer's instructions. Finally, the concentrations of the purified Snv4 and Snv6 were estimated using the Bradford assay with BSA as the standard according to standard protocols, and then the proteins were dispensed as 1 mL aliquots into 1.65 mL microcentrifuge tubes, flash frozen in liquid $N_2$, and stored at -80 °C until use.

*3.2.12. In vitro assays using Snv4 and Snv6*

To assay the function of Snv4 as a glycosyltransferase, an in vitro approach was taken based on a published study.[28] The 0.1 mL reactions were composed of 1 mM of a putative sugar donor, α-D-glucose-1-phosphate or UDP-glucuronic acid, 1 mM of a putative sugar acceptor, 2-aminopurine or guanine, in the Snv4 dialysis buffer described earlier supplemented with 1 mM $MgCl_2$ and were initiated by the addition of Snv4 to 10 μM. Reactions were incubated at 37 °C overnight and then filtered through 10 kDa molecular weight cutoff Nanosep® centrifugal filters by Pall Laboratory from VWR according to the manufacturer's instructions. Samples were analyzed by HPLC using the same equipment and methods as used earlier for isolation of the miharamycins. In alternative approaches, HPLC purified amipurimycin was reacted with 10 mM UDP in the presence of 10 μM Snv4 in the same reaction buffer used earlier or HPLC purified amipurimycin was incubated with 10 μM Snv4 in 50 mM or 100 mM potassium phosphate buffer (pH 8.0). These reactions were incubated overnight at 37 °C, filtered, and analyzed by HPLC as well. To assay the function of Snv6, 0.1 mg Snv6 was incubated with 0.5 mM NADH or NADPH and 0.5 mM guanine, guanosine, 5'-guanosine monophosphate, 2-deoxy 5'-guanosine monophosphate, guanosine diphosphate, or 2'-deoxy guanosine triphosphate in reaction buffer composed of 50 mM Tris HCl (pH 8.0), 0.5 mM DTT, and 10% glycerol. These reactions were incubated at 37 °C for 1.5 h and then filtered and analyzed as for the Snv4 reactions.

## 3.3. RESULTS AND DISCUSSION

*3.3.1. Isolation of the Miharamycins*

Our studies of the biosynthesis of miharamycins and amipurimycin actually began prior to obtaining the putative biosynthetic gene clusters and commenced with

development of a reliable method to producing and isolating these compounds. Interestingly, despite being reportedly deposited in the Institute of Fermentation Osaka (IFO) as IFO-13572,[29,30] Kuniko Sato, an IFO curator, revealed the deposit of *S. novoguineensis* was cancelled by the depositor. Therefore, we first turned our attention to *S. miharaensis*. The only available literature in English regarding the growth of S. miharaensis and the isolation of miharamycins was a US patent.[14] Frustratingly, our attempts to replicate the conditions described were met with failure. Finally, Dr. Masato Tani of Meiji Seika Kaisha, Ltd. provided copies of Meiji Seika Kenkyū Nenpō pertaining to the isolation and biological activity of the miharamycins,[15,21,31] and translation of these documents by Dr. Yasushi Ogasawara, a post-doctoral fellow in our lab, revealed the correct composition of the growth medium. However, according to the examples provided, the miharamycins were produced in large scale format, e.g., 300 L, using sparged air.[14,15] Because we could not replicate these conditions in our lab, we tried to optimize a small scale format to produce miharamycins. Various culture conditions were used, and a disc-diffusion bioassay against *P. aeruginosa* suggested in the literature from Meiji Seika Kaisha[21] was used to detect production of the miharamycins. We finally found the conditions described in Materials and Methods which reliably resulted in bioactivity by the disc-diffusion assay, but analysis by HPLC could not detect the miharamycins.

*3.3.2. Putative Precursor Feeding Studies*

First applied to the industrial production of penicillin G,[32,33] precursor feeding has been often used to increase yields of bacterial natural products.[33-36] We therefore undertook a program of putative precursor feeding to both increase production of miharamycins and give insight into the biosynthesis of these molecules. Of the compounds fed, L-arginine was first found to almost double the bioactivity of *S. miharaensis*

production culture extracts. This result was rationalized by considering the biosynthesis of arginine is metabolically "expensive" due to the consumption of two equivalents of ATP for each molecule of arginine produced.[37,38] Gratifyingly, the increased activity seemed to be related to increased production of miharamycins, as HPLC analysis was able to detect their presence in culture extracts. At supplementation of L-arginine to concentrations above 2 mM, the effects on the production of miharamycin were unpredictable, and at 20 mM, the production of miharamycins was apparently completely suppressed. However, supplementation with L-ornithine did not seem to have a ceiling to its effect on increasing miharamycins production. This finding is not too surprising, considering L-arginine regulation has pleitropic effects intracellularly.[38,39] The only other putative precursor found to increase miharamycins production was 5'-guanosine monophosphate (GMP), and the combination of GMP and L-ornithine qualitatively afforded synergistic increases in miharamycins production. The effect of GMP was somewhat surprising because nucleotides reportedly do not readily cross bacterial cell membranes, and it provides support for proposed origin of 2-aminopurine, as discussed in Chapter 2.

In order to validate our method of growth of S. miharaensis for miharamycins production, we performed a time-course study of miharamycins production using L-ornithine and 5'-GMP fed cultures, singly in combination with each other, at a concentration of 3 mM. Generally, concentrations of miharamycin B under each

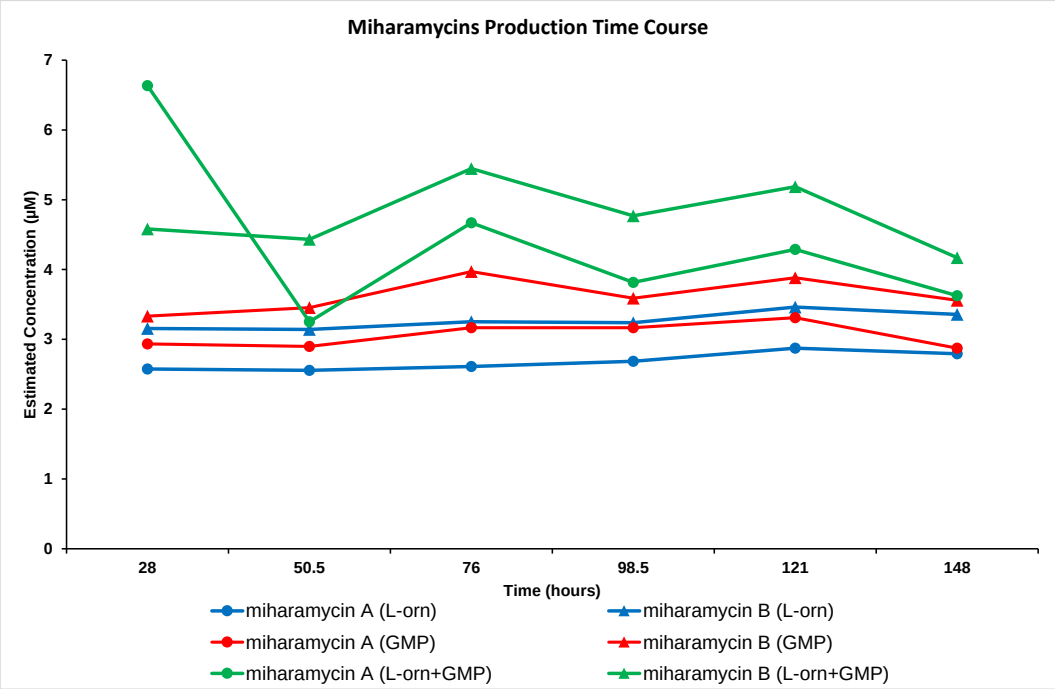



**Figure 3.6:** Miharamycins production time course

condition were higher than concentrations of miharamycin A, and these concentrations tended to stay steady during the course of the study. Notably, higher concentrations of miharamycins were obtained in the culture fed both L-ornithine and 5'-GMP, and these concentrations fluctuated over a larger range than the cultures fed only L-ornithine and 5'-GMP. Interestingly, the first time point at 28 h for the combination culture showed the highest concentration of miharamycin A observed during the study at 6.63 µM, but it seemed to decrease sharply by the second time point at 50.5 h. The concentration of miharamycin A did increase at later time points, but it never increased as high as was observed at 28 h. Unfortunately, the conclusions from this study can only be taken lightly considering the experiment was not replicated. However, it could reveal that future attempts to isolate miharamycin A would benefit not only from feeding both L-ornithine and 5'-GMP but also from ending fermentation soon after supplementation.

*3.3.3. Inhibitor Feeding Studies*

While putative precursor feeding has been used to provide insight in many cases, inspired by the work of Gould and co-workers,[40,41] we used a set of enzyme inhibitors in an effort to identify possible intermediates in the miharamycins biosynthetic pathway. These inhibitors included *O*-(carboxymethyl)-hydroxylamine (also known as aminooxyacetic acid), 6-diazo-5-oxo-L-norleucine, 1-aminobenzotriazole, and oxythiamine, which are inhibitors of PLP-dependent enzymes,[42,43] glutamine-utilizing enzymes,[44] cytochrome P450 enzymes,[45] and thiamine-pyrophosphate-dependent enzymes,[46] respectively. These specific inhibitors were used to: 1) try to isolate the miharamycins core saccharides by inhibiting installation of the amino functionality at C6 (aminooxy acetic acid and 6-diazo-5-oxo-L-norleucine); 2) determine the origin of the C8-C9 fragment (oxythiamine); and 3) determine if the hydroxylated arginine of miharamycin A is derived from a cytochrome P450-mediated process (1-aminobenzotriazole).

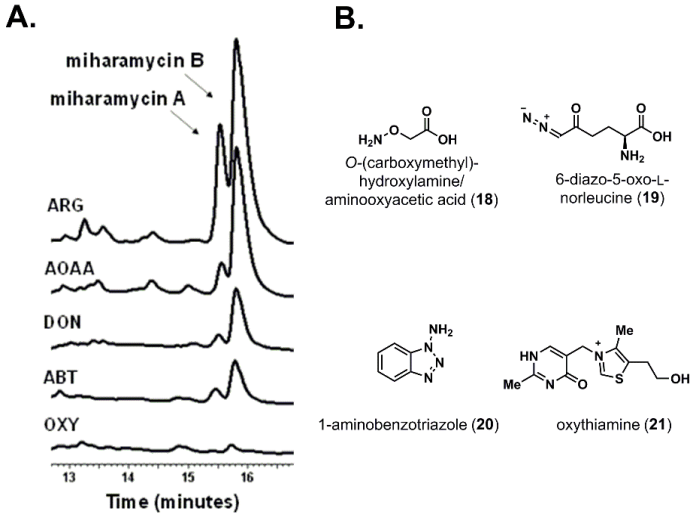



**Figure 3.7:** Inhibitor feedings and structures of inhibitors. **A.** HPLC analysis of inhibitor-fed cultures; **B.** Structures of inhibitors used: ARG= arginine, AOAA= aminooxyacetic acid, DON= 6-diazo-5-oxo-L-norleucine, ABT= 1-aminobenzotriazole, and OXY= oxythiamine

All inhibitors were fed at 2 mM except 6-diazo-5-oxo-L-norleucine, which was fed at 0.5 mM. Compared to a culture fed L-arginine at 2 mM, the aminooxyacetic acid fed culture produced a comparable amount of miharamycin B but considerably less miharamycin A; 6-diazo-5-oxo-L-norleucine and 1-aminobenzotriazole fed cultures produced less miharamycins in general; and the oxythiamine-fed culture produced almost no miharamycins. Because 6-diazo-5-oxo-L-norleucine and 1-aminobenzotriazole fed cultures seemed to have a similar effect, we believe these inhibitors may not have acted specifically on the biosynthesis of the miharamycins, making the interpretation of their effects less clear. It should be noted though that 6-diazo-5-oxo-L-norleucine appears to potently suppress miharamycins production, as it was only fed at 0.5 mM but attained the same level of product suppression as 2 mM 1-aminobenzotriazole. As discussed in Chapter 2, we believe 2-amino purine could result from guanine or guanosine, and considering 6-diazo-5-oxo-L-norleucine suppresses *de novo* purine biosynthesis through competitive inhibition of 2-*N*-formylglycinamide ribotide (FGAR) amidotransferase,[44] the suppression of miharamycins production could reflect decreased availability of purines, particularly guanosine, which would support our proposal for the origin of 2-aminopurine. On the other hand, aminooxyacetic acid and oxythiamine fed cultures led to more divergent effects on the production of the miharamycins.

Aminooxyacetic acid is a well-known inhibitor of aminotransferases, and it is thought to act through the formation of a stable oxime with pyridoxal.[42,43] In this light, the effect of aminooxyacetic acid on the production of miharamycin A cannot be explained, since the difference between miharamycins A and B is a hydroxyl group and not an amine. While aminooxyacetic acid is thought to act non-specifically,[47] very few reports of its activity on enzymes other than those dependent on PLP are available. However, aminooxyacetic acid has been observed to inhibit human FAD-dependent D-aspartate

oxidase[48] and spinach leaf NAD(P)H-dependent glyoxylate reductase.[49] In the former case, aminooxyacetic acid was characterized as a potent non-competitive inhibitor with no determination of its reversibility or lack thereof, and in the latter, aminooxyacetic acid was found to be a reversible inhibitor which preferentially bound the enzyme-glyoxylate complex. Nevertheless, the mechanism of aminooxyacetic acid inhibition of these enzymes was not investigated. Interestingly, aminooxyacetic acid was found to be able to form a stable complex with pentobarbital in vivo,[50] and though the mechanism was not investigated further, the only electrophilic functional group of this barbiturate is an amide.[51] Thus, it is plausible aminoxyacetic acid could inhibit the requisite enzyme(s) necessary for the oxidation of the arginine residue found in miharamycin A, albeit through an unknown mechanism. Furthermore, if this is indeed the case, the selective effect of aminooxyacetic acid on the production of the miharamycin A could indicate modification of arginine is a "tailoring" reaction performed on miharamycin B.

This line of reasoning is also supported by the YcgI/YcgG homologue found in the *smh* cluster, Smh24. As discussed in Chapter 2, even though the phosphonate natural product argolaphos contains the same hydroxylated arginine,[52] the YcgI/YcgG homologue from its biosynthetic gene cluster displays only modest similarity and identity to Smh24, which suggests their divergent primary sequences are driven by the dissimilarity of their substrates. According to sequence analyses presented in Chapter 2, Smh24 contains a putative Rossmann fold, which could be involved in binding a flavin cofactor, supporting the role of this enzyme as a putative oxidoreductase.

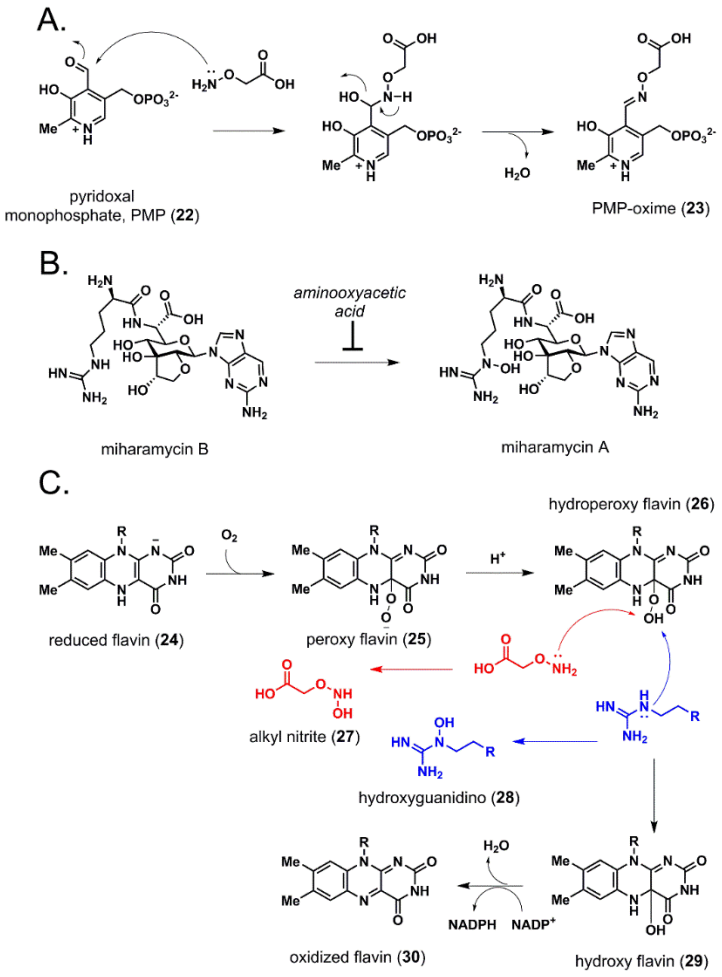



**Figure 3.8: A.** mechanism of aminooxyacetic acid inhibition of PMP; **B.** presumed effect of feeding aminooxyacetic acid; **C.** hypothetical explanation for observed suppression of miharamycin A production in the presence of aminooxyacetic acid

In this light, and given the reactivity of aminooxyaceitc acid as a nucleophile, the selective effect of this inhibitor on the production of miharamycin A can be rationalized as follows. Flavin-dependent oxidoreductases are well-known for their ability to perform heteroatom oxidation reactions, such as *N*-oxidation. Usually, hydroxylation is mediated through a reactive species, such as peroxy or hydroperoxy flavin. The former is implicated in electrophilic oxygenation reactions, whereas the latter is implicated in nucleophilic

oxygenation reactions.[53,54] As aminooxyacetic acid is a reactive nucleophile, it could be envisioned to more readily react with hydroperoxy flavin, resulting in the formation of an alkyl nitrite. This sequence implies aminooxyacetic acid would be a competitive inhibitor preventing oxidation of arginine or miharamycin B to the corresponding compounds hydroxyarginine or miharamycin A, respectively. More work is needed to explore this possibility.

Oxythiamin has been characterized as an inhibitor of TPP-dependent enzymes, particularly transketolase. Although its mechanisms of action are poorly characterized, it can be pyrophosphorylated in vivo to generate a non-functional analog which can displace TPP.[46] As discussed in Chapter 2, the miharamycins are built upon a perhydrofuropyran, but unlike the octosyl acid derivatives, we felt the furan portion was derived from a cyclized two-carbon branch, especially given the structure of amipurimycin,[55] whose C8-C9 fragment could be viewed as an "open chain" form of the furan ring found in the miharamycins. Because transketolase-like enzymes have been implicated in the biosynthesis of two-carbon branched sugars like the eurekanate moiety of avilamycin and everninomicin,[56,57] we felt oxythiamine was a good choice to explore this possibility. Instead of recovering intermediates lacking the two-carbon branch, oxythiamine all but ablated production of the miharamycins, indicating a TPP-dependent enzyme plays an important role in the biosynthesis of the miharamycins or at least is responsible for a step prior to and necessary for the attachment of the 2-aminopurine chromophore. The role of a transketolase in the biosynthesis of the miharamycins and amipurimycin was substantiated by the identification of transketolases in the smh and snv clusters, Smh5a/b and Snv3a/b, respectively. In consideration of the predicted function of Smh5a/b and Snv3a/b in the installation of the two-carbon branches, and the effect of oxythiamine on the production of

the miharamycins, branching is likely a required step before 2-aminopurine can be attached.

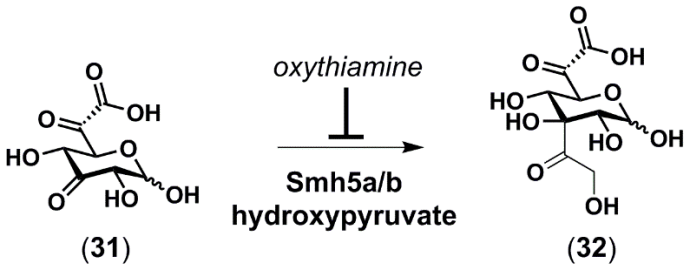



**Figure 3.9:** Presumed effect of oxythiamine feeding on miharamycins production

Later, assistance from Hisayuki Komaki, a curator from the NITE Biological Resource Center (NBRC) in Japan, led to the location of *S. novoguineensis* in the CBS-KNAW culture collection in the Netherlands. Despite the high structural similarity between amipurimycin and the miharamycins, the available literature for the growth of *S. novoguineensis* and isolation of amipurimycin described vastly different approaches from those described for *S. miharaensis* and the miharamycins. Early attempts to produce and isolate amipurimycin following the published procedure[29,30] were unsuccessful, so the protocol we developed for the production and isolation of miharamycins was utilized. Remarkably, this approach led to the isolation of amipurimycin. Unlike *S. miharaensis*, *S. novoguineensis* did not respond to feeding of 5'-GMP. Prior to identification of the amipurimycin biosynthetic gene cluster, which, as discussed in Chapter 2, indicates the pendant amino acid cispentacin is derived from glutarate semialdehyde, we believed this unusual amino acid could be derived from cyclization of L-lysine or L-norleucine, but feeding of these amino acids did not lead to increased production of amipurimycin.

*3.3.4. In vitro Assays with Snv4 and Snv6*

Then, we sought to determine the early steps of amipurimycin biosynthesis through the use of *in vitro* assays. The biosyntheses of the hPNAs such as blasticidin S, mildiomycin, and gougerotin commence with the condensation of UDP-glucuronic acid and cytosine,[9] and though the content of the *smh* and *snv* clusters identified in Chapter 2 imply a PKS-based biosynthetic pathway for amipurimycin and the miharamycins, it was important to rule out the involvement of a biosynthetic pathway more akin to the known hPNAs. To this end, N-terminally $His_6$-tagged Snv4 was expressed and purified. Under the conditions assayed, Snv4 was not able to condense 2-aminopurine or guanine to UDP-glucuronic acid. Alternatively, a reverse glycosylation assay using UDP was not able to remove the 2-aminopurine base from HPLC purified amipurimycin. In consideration of the pathway discussed in Chapter 2, wherein the core nucleosides of the miharamycins and amipurimycin are activated as monophosphates for transfer to guanine or 2-aminopurine, a reverse glycosylation assay in phosphate buffer to remove 2-aminopurine from amipurimycin was attempted, as well as a forward glycosylation assay to glycosylate guanine or 2-amino purine with the core nucleoside mimic glucose-1-phosphate. However, these attempts were not successful.

In consideration of the stimulatory effect 5'-GMP had on the production of the miharamycins, guanine or its nucleoside or nucleotides were thought to be potential precursors to 2-aminopurine through direct reduction of the C6 carbonyl, as discussed in Chapter 2. Therefore, we cloned and expressed Snv6 as an N-terminally $His_6$-tagged protein.

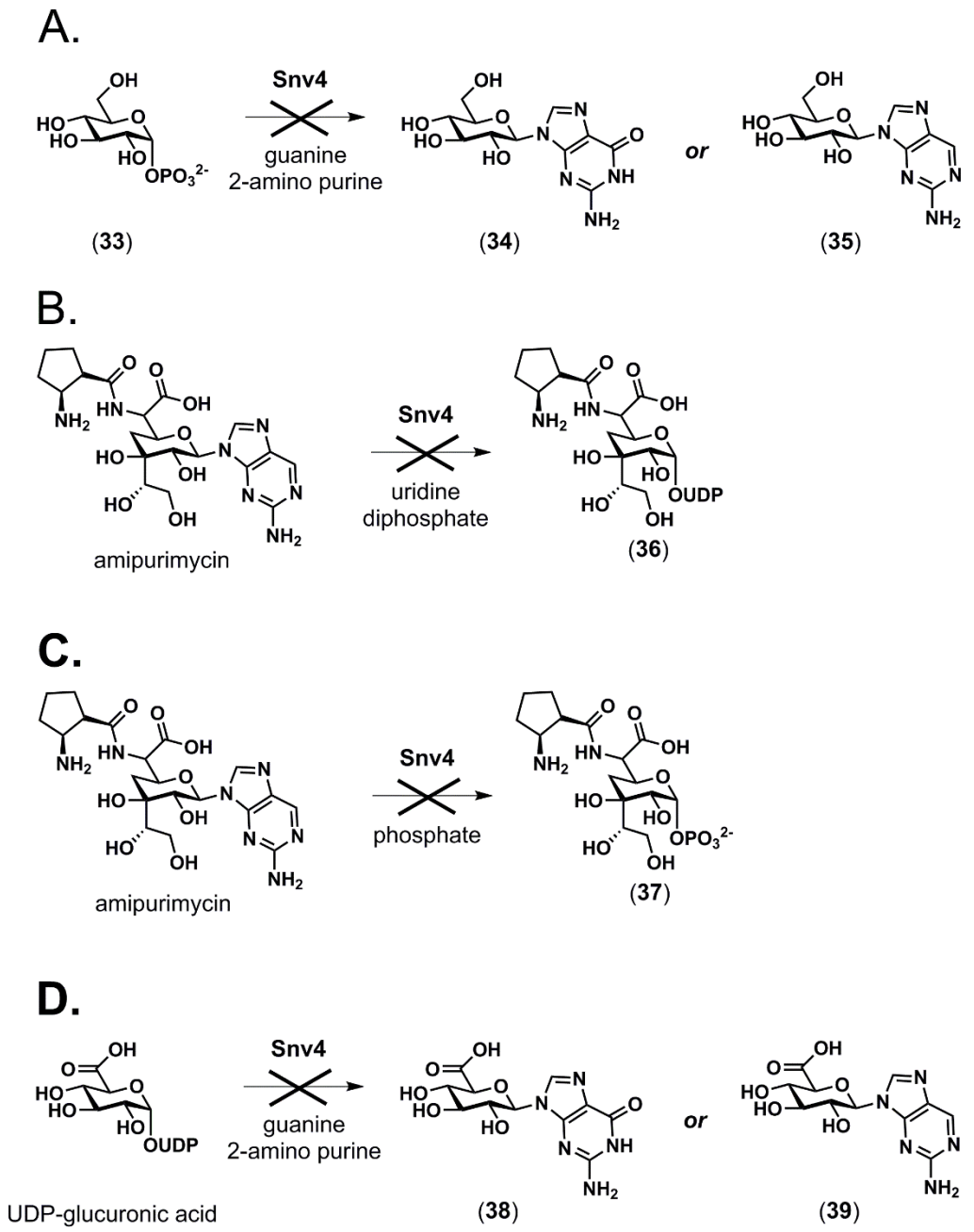



**Figure 3.10:** Summary of Snv4 assays performed.

Interestingly, the purified protein was a deep brown/orange color, which suggested the presence of a chromogenic cofactor. We discussed in Chapter 2 the possibility Snv6

binds a cofactor based on the presence of a putative Rossmann motif[58] as well as the possibility the protein binds iron due to the presence of a cysteine-rich motif typical of iron-sulfur cluster-binding proteins.[59] As expected, analysis of recombinant Snv6 by HPLC spectroscopic methods by Mr. Geng-Min Lin, a graduate student in the Liu Lab, revealed the presence of one equivalent of FAD per molecule and, preliminarily, supported the existence of a 4Fe-4S cluster (data not shown). The ability of Snv6 to reduce guanine, guanosine, 5'-GMP, 2-deoxy 5'-GMP, GDP, or 2-deoxy GTP using NADH or NADPH was assayed under aerobic conditions and, by Mr. Geng-Min Lin, under anaerobic conditions, but no reaction was observed. It is possible the reaction is more complex, requiring additional cofactors or components, or 2-aminopurine is not derived from guanine or its derivatives.

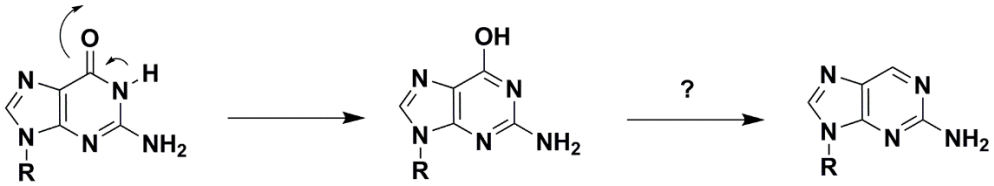



**Figure 3.11:** Hypothetical "direct" reduction of guanine or a derivative thereof by Snv6 to lead to 2-aminopurine

*3.3.5. Genetic Manipulation of S. novoguineensis*

Next, we sought to provide more conclusive evidence of the involvement of the *smh* and *snv* clusters in the biosynthesis of the miharamycins and amipurimycin through gene inactivation experiments. *S. miharaensis* was recalcitrant towards attempts to introduce DNA, and these attempts were further hindered by the inability of S. miharaensis to reliably sporulate. In fact, lack of sporulation except under very specific conditions was

reported to be a defining characteristic of this strain. Although this phenomenon was not explored in *S. miharaensis*, in consideration of the effect of exogenously supplied arginine on miharamycins production, it is notable that defects in sporulation have been linked to unstable arginine biosynthesis genes in *Streptomyces lividans 66*.[60] On the other hand, *S. novoguineensis* sporulated abundantly and quite easily on several solid media, which made intergeneric conjugation an attractive approach to gene inactivation in this strain, as this method can even circumvent active restriction endonuclease systems.[12] Furthermore, we felt the apparently constitutive production of amipurimycin would be a more convenient phenotype. Single crossover mutations are notable for their ease of introduction,[12] so we selected this method to interrupt *snv8*. This gene was chose for two reasons. First of all, since it is a large gene (3003 bp) a larger fragment could be utilized for the integration plasmid, which could improve efficiency of integration.[12] Secondly, by interrupting this PKS-encoding gene, we could obtain some evidence for its involvement in the production of amipurimycin.

The plasmid pKCsnv8Int was easily introduced into *S. novoguineensis* through intergeneric conjugation via *E. coli* S17-1, a methylase proficient strain,[23,24] which indicates *S. novoguineensis* may not have a high restriction barrier.[61] Interestingly, in the course of introducing plasmid pKCsnvInt8 into *S. novoguineensis*, we found this strain displays basal resistance to apramycin, but this could be overcome by using a concentration of this antibiotic of at least 100 μg/mL in all liquid and solid media. After three rounds of selection, three putative integrants, 94-A, 93-3, and 99-A were recovered. Because integration of the plasmid pKCsnv8Int would result in a ~7.5 kbp insertion in *snv8*, PCR amplification of interrupted *snv8* (~10.5 kbp) was expected to be inefficient under conditions optimized to amplify the native ~3 kbp gene. Compared to the wild-type strain of *S. novoguineensis*, this PCR screening failed 94-A, 93-3, and 99-A, indicating the three

recovered integrants did indeed harbor an interrupted copy of *snv8*. Strain 94-A was assayed for its ability to produce amipurimycin compared to the wild-type strain, and no amipurimycin was detected. Therefore, we are confident the *snv* cluster truly is the amipurimycin biosynthetic gene cluster and, because of its high homology to the *smh* cluster we are confident the miharamycins biosynthetic gene cluster has been identified. Finally, to further aid in future gene deletion studies through λRED-mediated recombineering and heterologous expression, we utilized cosmid library screening to isolate cosmid pSNVNP2B47 which appears to contain the entire amipurimycin biosynthetic gene cluster.

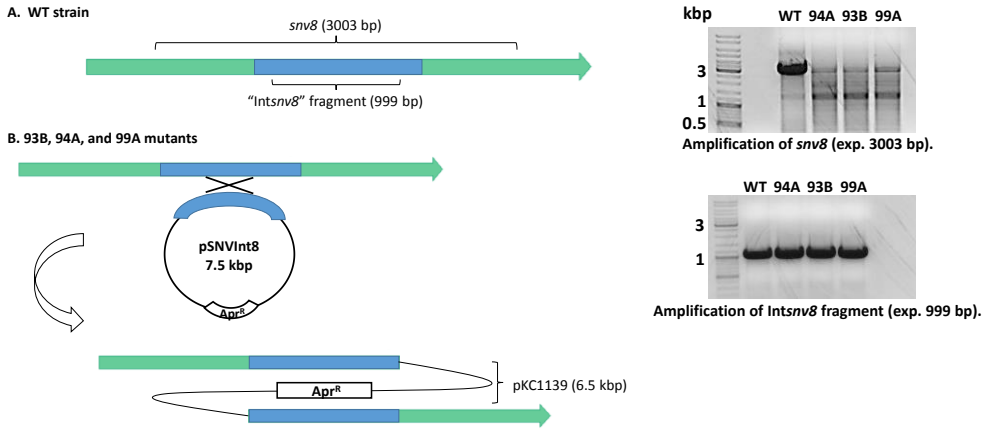



**Figure 3.12:** Schematic of snv8 single crossover interruption and PCR screening of exconjugants

## 3.4. CONCLUSIONS

While the fine details of the biosynthetic pathways of the miharamycins and amipurimycin are currently lacking, our work provides some insight into the biosyntheses of these complex PNA. Specifically, we have developed reliable methods for the production and isolation of the miharamycins and amipuimycin, and, most importantly, we have provided the first evidence that the clusters identified in Chapter 2 are indeed the

correct clusters through single crossover mutation of *snv8*. Furthermore, this work has established a method for the introduction of DNA into *S. novoguineensis* and led to the isolation of the amipurimycin gene cluster on a single cosmid, a potentially useful tool in future investigations into the biosynthesis of amipurimycin.

Although the predicted activities of recombinant Snv4 and Snv6 were not observed under the conditions used, some important conclusions may be drawn. Namely, in support of the PKS model elaborated in Chapter 2 and unlike the biosynthesis of other hPNA, the biosyntheses of amipurimycin and miharamycin do not likely involve the condensation of 2-aminopurine or guanine to UDP-glucuronic acid. Furthermore, even though 5'-GMP was observed to be a strong inducer of miharamycins biosynthesis, which implies it could be involved in the biosynthesis of miharamycins and amipurimycin, the inactivity of Snv6 towards guanine or its derivatives suggests 2-aminopurine may be formed after construction of the core nucleoside. On the other hand, it is very possible the reaction of Snv6 is more complex than our conditions, possibly requiring other cofactors or partner enzymes, or the function of Snv6 could be sensitive to the presence of a $His_6$-tag. Furthermore, if Snv6 does indeed harbor an iron-sulfur cluster, due to the lability of this cofactor under aerobic conditions, it would need to be reconstituted anaerobically before use. In all, several avenues to investigate the putative reaction are available.

The work presented in this chapter dovetails with the analysis provided in Chapter 2, thereby setting the stage for deeper study of the biosyntheses of the miharamycins and amipurimycin, which could be the first members of PKS-derived peptidyl nucleosides and offer insight into rare examples of novel enzymatic catalysis.

## 3.5. REFERENCES


(1) Kim, E.; Moore, B. S.; Yoon, Y. J. Reinvigorating natural product combinatorial biosynthesis with synthetic biology. *Nat Chem Biol.* **2015**, *11*, 649.

(2) Rutledge, P. J.; Challis, G. L. Discovery of microbial natural products by activation of silent biosynthetic gene clusters. *Nat Rev Microbiol.* **2015**, *13*, 509.

(3) Zarins-Tutt, J. S.; Barberi, T. T.; Gao, H.; Mearns-Spragg, A.; Zhang, L.; Newman, D. J.; Goss, R. J. Prospecting for new bacterial metabolites: a glossary of approaches for inducing, activating and upregulating the biosynthesis of bacterial cryptic or silent natural products. *Nat Prod Rep.* **2015**.

(4) Cone, M. C.; Yin, X.; Grochowski, L. L.; Parker, M. R.; Zabriskie, T. M. The blasticidin S biosynthesis gene cluster from Streptomyces griseochromogenes: sequence analysis, organization, and initial characterization. *Chembiochem.* **2003**, *4*, 821.

(5) Wu, J.; Li, L.; Deng, Z.; Zabriskie, T. M.; He, X. Analysis of the mildiomycin biosynthesis gene cluster in Streptoverticillum remofaciens ZJU5119 and characterization of MilC, a hydroxymethyl cytosyl-glucuronic acid synthase. *Chembiochem.* **2012**, *13*, 1613.

(6) Feng, J.; Wu, J.; Gao, J.; Xia, Z.; Deng, Z.; He, X. Biosynthesis of the beta-methylarginine residue of peptidyl nucleoside arginomycin in Streptomyces arginensis NRRL 15941. *Appl Environ Microbiol.* **2014**, *80*, 5021.

(7) Niu, G.; Li, L.; Wei, J.; Tan, H. Cloning, heterologous expression, and characterization of the gene cluster required for gougerotin biosynthesis. *Chem Biol.* **2013**, *20*, 34.

(8) Zhang, G.; Zhang, H.; Li, S.; Xiao, J.; Zhu, Y.; Niu, S.; Ju, J.; Zhang, C. Characterization of the amicetin biosynthesis gene cluster from Streptomyces vinaceusdrappus NRRL 2363 implicates two alternative strategies for amide bond formation. *Appl Environ Microbiol.* **2012**, *78*, 2393.

(9) Niu, G.; Tan, H. Nucleoside antibiotics: biosynthesis, regulation, and biotechnology. *Trends Microbiol.* **2015**, *23*, 110.

(10) Walsh, C. T.; Zhang, W. Chemical logic and enzymatic machinery for biological assembly of peptidyl nucleoside antibiotics. *ACS Chem Biol.* **2011**, *6*, 1000.

(11) Sambrook, J.; Russell, D. W. *Molecular cloning : a laboratory manual*; Third ed.; Cold Spring Harbor Laboratory: Cold Spring Harbor, NY, 2001.

(12) Kieser, T.; Bibb, M. J.; Buttner, M. J.; Chater, K. F.; Hopwood, D. A. *Practical streptomyces genetics*; The John Innes Foundation: Norwich, U.K., 2000.

(13) QIAgen Miniprep Handbook. Fourth ed. 2015.

(14) Niida, T.; Yumoto, H.; Tsuruoka, T.; Hamamoto, K.; Shomura, J.; Ohashi, T.; Antibiotics Obtained from Streptomyces Miharaensis; U.S. Patent 3,678,159, July 18, 1972.

(15) Tsuruoka, T.; Yumoto, H.; Ezaki, N.; Niida, T. New antibiotics, miharamycins A and B. I. Isolation and characterization of miharamycins A and B. *Meiji Seika Kenkyu Nenpo.* **1967**, *9*, 1.

(16) Corman, J.; Tsuchiya, H. M.; Koepsell, H. J.; Benedict, R. G.; Kelley, S. E.; Feger, V. H.; Dworschack, R. G.; Jackson, R. W. Oxygen absorption rates in laboratory and pilot plant equipment. *Appl Microbiol.* **1957**, *5*, 313.

(17) McDaniel, L. E.; Bailey, E. G. Effect of shaking speed and type of closure on shake flask cultures. *Appl Microbiol.* **1969**, *17*, 286.

(18) Brown, E. G.; Konuk, M. Biosynthesis of nebularine (purine 9-B-D-ribofuranoside) involves enzymic release of hydroxylamine from adenosine. *Phytochemistry.* **1995**, *38*, 61.

(19) Jorpes, E.; Thoren, S. The use of the Sakaguchi reaction for the quantitative determination of arginine. *Biochem J.* **1932**, *26*, 1504.

(20) Bhattacharya, K. R.; Datta, J.; Roy, D. K. A study on the Sakaguchi reaction of arginine on filter paper. *Arch Biochem Biophys.* **1958**, *77*, 297.

(21) Shomura, T.; Hamamoto, K.; Ohashi, T.; Amano, S.; Yoshida, J.; Moriyama, C.; Niida, T. New antibiotics, miharamycins A and B. II. Some biological characteristics of miharamycin. *Meiji Seika Kenkyu Nenpo.* **1967**, *9*, 5.

(22) Biermann, M.; Logan, R.; O'Brien, K.; Seno, E. T.; Nagaraja, R. R.; Schoner, B. E. Plasmid cloning vectors for the conjugal transfer of DNA from Eschericia coli to Streptomyces spp. *Gene.* **1992**, *116*, 43.

(23) Simon, R.; Priefer, U.; Puhler, A. A broad host range mobilization system for in vivo genetic engineering: transposon mutagenesis in Gram negative bacteria. *Nature Biotechnology.* **1983**, *1*, 784.

(24) Mazodier, P.; Petter, R.; Thompson, C. Intergeneric conjugation between Escherichia coli and Streptomyces species. *J Bacteriol.* **1989**, *171*, 3583.

(25) Baltz, R. H.; Matsushima, P. Protoplast fusion in Streptomyces: conditions for efficient genetic recombination and cell regeneration. *J Gen Microbiol.* **1981**, *127*, 137.

(26) The pET Expression System. Eleventh ed. 2006.

(27) The QIAexpressionist. Fifth ed. 2003.

(28) Gould, S. J.; Guo, J. Cytosylglucuronic acid synthase (cytosine: UDP-glucuronosyltransferase) from Streptomyces griseochromogenes, the first prokaryotic UDP-glucuronosyltransferase. *J Bacteriol.* **1994**, *176*, 1282.

(29) Harada, S.; Kishi, T. Isolation and characterization of a new nucleoside antibiotic, amipurimycin. *J Antibiot (Tokyo).* **1977**, *30*, 11.

(30) Iwasa, T.; Kishi, T.; Matsuura, K.; Wakae, O. Streptomyces novoguineensis sp. Nov., an amipurimycin producer, and antimicrobial activity of amipurimycin. *J Antibiot (Tokyo).* **1977**, *30*, 1.

(31) Noguchi, T.; Komoto, K.; Yasuda, Y.; Hashimoto, S.; Niida, T. New antibiotics, miharamycins A and B. III. Control activity of miharamycin against plant disease. *Meiji Seika Kenkyu Nenpo.* **1967**, *9*, 11.

(32) Soltero, F. V.; Johnson, M. J. The effect of the carbohydrate nutrition on penicillin production by Penicillium chrysogenum Q-176. *Appl Microbiol.* **1953**, *1*, 52.

(33) Elander, R. P. Industrial production of beta-lactam antibiotics. *Appl Microbiol Biotechnol.* **2003**, *61*, 385.

(34) Jiang, L.; Wei, J.; Li, L.; Niu, G.; Tan, H. Combined gene cluster engineering and precursor feeding to improve gougerotin production in Streptomyces graminearus. *Appl Microbiol Biotechnol.* **2013**, *97*, 10469.

(35) Du, W.; Huang, D.; Xia, M.; Wen, J.; Huang, M. Improved FK506 production by the precursors and product-tolerant mutant of Streptomyces tsukubaensis based on genome shuffling and dynamic fed-batch strategies. *J Ind Microbiol Biotechnol.* **2014**, *41*, 1131.

(36) Reeves, A. R.; Brikun, I. A.; Cernota, W. H.; Leach, B. I.; Gonzalez, M. C.; Weber, J. M. Effects of methylmalonyl-CoA mutase gene knockouts on erythromycin production in carbohydrate-based and oil-based fermentations of Saccharopolyspora erythraea. *Journal of Industrial Microbiology and Biotechnology.* **2006**, *33*, 600.

(37) Caldara, M.; Dupont, G.; Leroy, F.; Goldbeter, A.; De Vuyst, L.; Cunin, R. Arginine biosynthesis in Escherichia coli: experimental perturbation and mathematical modeling. *J Biol Chem.* **2008**, *283*, 6347.

(38) Rodriguez-Garcia, A.; Ludovice, M.; Martin, J. F.; Liras, P. Arginine boxes and the argR gene in Streptomyces clavuligerus: evidence for a clear regulation of the arginine pathway. *Mol Microbiol.* **1997**, *25*, 219.

(39) Makarova, K. S.; Mironov, A. A.; Gelfand, M. S. Conservation of the binding site for the arginine repressor in all bacterial lineages. *Genome Biol.* **2001**, *2*, RESEARCH0013.

(40) Gould, S. J.; Guo, J.; Geitmann, A.; DeJesus, K. Nucleoside intermediates in blasticidin S biosynthesis identified by the in vivo use of enzyme inhibitors. *Canadian Journal of Chemistry.* **1994**, *72*, 6.

(41) Zhang, Q.; Gould, S. J.; Zabriskie, T. M. A new cytosine glycoside from Streptomyces griseochromogenes produced by the use in vivo of enzyme inhibitors. *J Nat Prod.* **1998**, *61*, 648.

(42) Liu, W.; Peterson, P. E.; Carter, R. J.; Zhou, X.; Langston, J. A.; Fisher, A. J.; Toney, M. D. Crystal structures of unbound and aminooxyacetate-bound Escherichia coli gamma-aminobutyrate aminotransferase. *Biochemistry.* **2004**, *43*, 10896.

(43) Loscher, W.; Honack, D.; Gramer, M. Use of inhibitors of gamma-aminobutyric acid (GABA) transaminase for the estimation of GABA turnover in various brain regions of rats: a reevaluation of aminooxyacetic acid. *J Neurochem.* **1989**, *53*, 1737.

(44) Kisner, D. L.; Catane, R.; Muggia, F. M. In *Cancer Chemo- and Immunopharmacology. 1. Chemopharmacology*; Mathe, G., Muggia, F. M., Eds.; Springer Berlin Heidelberg: Berlin, 1980; Vol. 74, p 258.

(45) Ortiz de Montellano, P. R.; Mathews, J. M.; Langry, K. C. Autocatalytic inactivation of cytochrome p-450 and chloroperoxidase by 1-aminobenzotriazole and other aryne precursors. *Tetrahedron* **1984**, *40*, 511.

(46) Wang, J.; Zhang, X.; Ma, D.; Lee, W. N.; Xiao, J.; Zhao, Y.; Go, V. L.; Wang, Q.; Yen, Y.; Recker, R.; Xiao, G. G. Inhibition of transketolase by oxythiamine altered dynamics of protein signals in pancreatic cancer cells. *Exp Hematol Oncol.* **2013**, *2*, 18.

(47) Kleczkowski, L. A. Inhibitors of photosynthetic enzymes/ carriers and metabolism. *Annual Review of Plant Physiology and Plant Molecular Biology.* **1994**, *45*, 339.

(48) Katane, M.; Kawata, T.; Nakayama, K.; Saitoh, Y.; Kaneko, Y.; Matsuda, S.; Miyamoto, T.; Sekine, M.; Homma, H. Characterization of the enzymatic and structural properties of human D-aspartate oxidase and comparison with those of the rat and mouse enzymes. *Biol Pharm Bull.* **2015**, *38*, 298.

(49) Kleczkowski, L. A.; Randall, D. D.; Blevins, D. G. Inhibition of Spinach Leaf NADPH(NADH)-Glyoxylate Reductase by Acetohydroxamate, Aminooxyacetate, and Glycidate. *Plant Physiol.* **1987**, *84*, 619.

(50) DaVanzo, J. P.; Matthews, R. J.; Stafford, J. E. Studies on the mechanism of action of aminooxyacetic acid. I. Reversal of aminooxyacetic acid-induced convulsions by various agents. *Toxicology and Applied Pharmacology.* **1964**, *6*, 388.

(51) In *PubChem Compound Database*; National Center for Biotechnology Information. .

(52) Ju, K. S.; Gao, J.; Doroghazi, J. R.; Wang, K. K.; Thibodeaux, C. J.; Li, S.; Metzger, E.; Fudala, J.; Su, J.; Zhang, J. K.; Lee, J.; Cioni, J. P.; Evans, B. S.; Hirota, R.; Labeda, D. P.; van der Donk, W. A.; Metcalf, W. W. Discovery of phosphonic acid natural products by mining the genomes of 10,000 actinomycetes. *Proc Natl Acad Sci U S A.* **2015**, *112*, 12175.

(53) van Berkel, W. J.; Kamerbeek, N. M.; Fraaije, M. W. Flavoprotein monooxygenases, a diverse class of oxidative biocatalysts. *J Biotechnol.* **2006**, *124*, 670.

(54) Mansoorabadi, S. O.; Thibodeaux, C. J.; Liu, H. W. The diverse roles of flavin coenzymes--nature's most versatile thespians. *J Org Chem.* **2007**, *72*, 6329.

(55) Goto, T.; Toya, Y.; Kondo, T. Structure of amipurimycin, a new nucleoside antibiotic produced by Streptomyces novoguineensis. *Nucleic Acids Symp Ser.* **1980**, s73.

(56) Treede, I.; Hauser, G.; Muhlenweg, A.; Hofmann, C.; Schmidt, M.; Weitnauer, G.; Glaser, S.; Bechthold, A. Genes involved in formation and attachment of a two-carbon chain as a component of eurekanate, a branched-chain sugar moiety of avilamycin A. *Appl Environ Microbiol.* **2005**, *71*, 400.

(57) Hofmann, C.; Boll, R.; Heitmann, B.; Hauser, G.; Durr, C.; Frerich, A.; Weitnauer, G.; Glaser, S. J.; Bechthold, A. Genes encoding enzymes responsible for biosynthesis of L-lyxose and attachment of eurekanate during avilamycin biosynthesis. *Chem Biol.* **2005**, *12*, 1137.

(58) Kleiger, G.; Eisenberg, D. GXXXG and GXXXA motifs stabilize FAD and NAD(P)-binding Rossmann folds through C(alpha)-H... O hydrogen bonds and van der waals interactions. *J Mol Biol.* **2002**, *323*, 69.

(59) Layer, G.; Gaddam, S. A.; Ayala-Castro, C. N.; Ollagnier-de Choudens, S.; Lascoux, D.; Fontecave, M.; Outten, F. W. SufE transfers sulfur from SufS to SufB for iron-sulfur cluster assembly. *J Biol Chem.* **2007**, *282*, 13342.

(60) Altenbuchner, J.; Cullum, J. DNA amplification and an unstable arginine gene in Streptomyces lividans 66. *Mol Gen Genet.* **1984**, *195*, 134.

(61) MacNeil, D. J.; Gewain, K. M.; Ruby, C. L.; Dezeny, G.; Gibbons, P. H.; MacNeil, T. Analysis of Streptomyces avermitilis genes required for avermectin biosynthesis utilizing a novel integration vector. *Gene.* **1992**, *111*, 61.

# Chapter 4: Investigation Into the Biosynthesis of TDP-2-Deoxy Nogalamine

## 4.1. INTRODUCTION

Nogalamycin (**1**) is a structurally unique anthracycline produced by *Streptomyces nogalater* NRRL 3035. Its high toxicity precludes its use in humans (Figure 4.1).[1,2] Nogalamycin belongs to a small group of anthracycline natural products characterized by an algycon extended by two additional heterocycles, rings E and F.[3,4] Recent investigations into the biosynthesis of nogalamycin have revealed that ring F is actually derived from a deoxy amino sugar, nogalamine (**2**), and ring E is assembled through two bond formation between nogalamine and ring D of the algycon.[1,5]

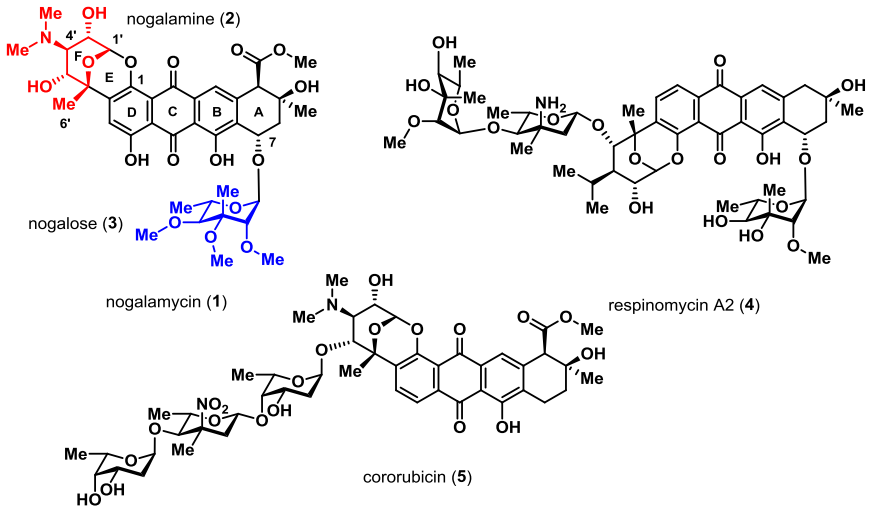



**Figure 4.1:** Sturctures of nogalamycin and representative extended ring system anthracyclines

One of these bonds is an *O*-glycosidic bond between nogalamine and the C1-hydroxyl of ring D, but the other is a rarely encountered *C*-glycosidic linkage between C5' of nogalamine and C2 of ring D.

Glycosylation at ring D of the anthracycline distinguishes nogalamycin and other related anthracycline systems such as respinomycin A2 (**4**) and cororubicin (**5**) from the prototypical, clinically useful anthracyclines daunorubicin (**6**) and doxorubicin (**7**) and their derivatives (**8**-**9**) which bear a single amino sugar attached to the C7-hydroxyl.[1,6,7] Interestingly, nogalamycin features the permethylated neutral deoxy sugar nogalose (**3**) at its C7-hydroxyl.[1]

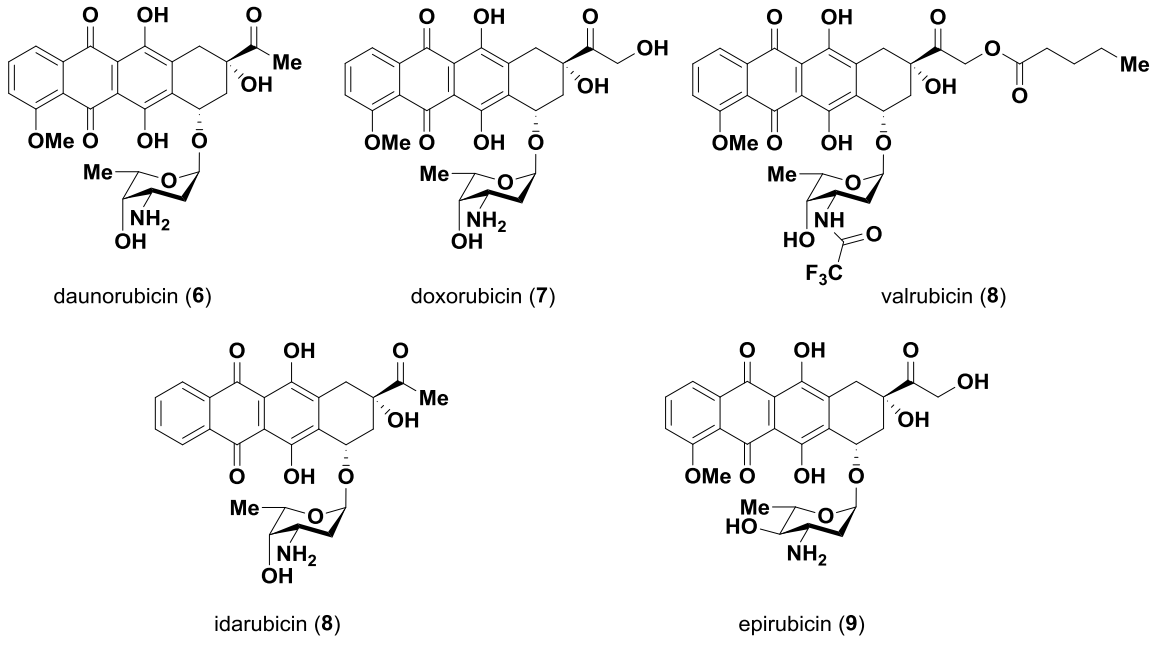



**Figure 4.2:** Structures of clinically relevant anthracyclines.

Similar to many other anthracyclines, nogalamycin is a potent DNA intercalator. Its dually glycosylated aglycon affords this natural product a distinct mode of DNA interaction, where the appended sugars contribute both minor and major groove interactions to stabilize the anthracycline:DNA complex.[6,8] Motivated by the wide clinical utility of daunorucin and doxorubicin, an extensive semisynthesis program to find less toxic nogalamycin derivatives was initiated. This effort yielded, among other molecules, menogaril, a 7-methoxy derivative of nogalamycin, which was found to be less toxic than nogalamycin but with low activity in clinical trials for a variety of cancers.[9-11] While semisynthetic efforts to modify nogalamycin mainly focused on derivatives lacking nogalose or other ring A substituents, little was done to alter the D-ring or its amino sugar, perhaps due to the stability endowed to this portion of the molecule by the additional carbon-carbon bond between nogalamine and the aglycon.[9,12,13]

Considering that the amino sugars decorating anthracyclines are usually important determinants of their potency, it is remarkable that little work has specifically focused the biosynthesis of these moieties.[3] Nogalamine is an ideal amino sugar to study because its biosynthetic pathway has not been definitively established. Bacterial deoxy sugars have been extensively shown to be derived from highly conserved biosynthetic pathways via NDP-activated (usually TDP) intermediates,[14] and analysis of the nogalamycin biosynthetic gene cluster does allow the proposal of a cursory pathway for nogalamine biosynthesis (Figure 4.3).[5] Namely, following established biosynthetic logic, SnogJ, a putative thymidylyl transferase; SnogK, a putative 4',6'-dehydratase; and SnogH, a

putative 2',3'-dehydratase can be envisioned to perform the first three conserved steps in deoxy sugar biosynthesis.

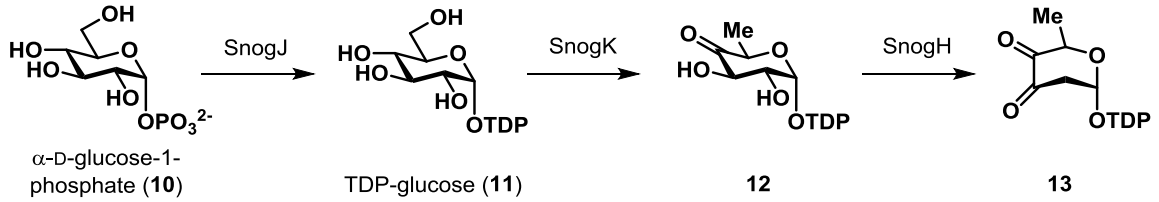



**Figure 4.3:** Pathway supporting a C2'-deoxy sugar pathway for nogalamine

Intriguingly, nogalamine, as found attached to the anthracycline, bears a hydroxyl at C2' which would suggest a pathway that keeps this hydroxyl intact, but the presence of a 2',3'-dehydratase and the lack of a 3',4'-keto isomerase, e.g., Tyl1a,[15] suggest that nogalamine follows a 2'-deoxy sugar biosynthetic pathway and that the C2' hydroxyl is installed at some point after it is removed, hence the proposed first steps above.[1,5] Currently, the 2' hydroxyl of nogalamine is proposed to be installed later in the pathway, perhaps after nogalamine is attached to the anthracycline, since noglamycin analogs lacking this substituent have been isolated from culture filtrates of *S. nogalater*.[5] There remain some significant ambiguities in the next steps. First, based on the presence of only one putative sugar epimerase in the nogalamycin biosynthetic gene cluster, SnogF, TDP-nogalamine and TDP-nogalose are likely formed through L-configured intermediates, and their pathways, therefore, must share this epimerase.[3,5] Second, the nogalamycin gene cluster has two sugar 4'-keto reductases, SnogG and SnogC; while SnogG is proposed to be

involved in nogalamine biosynthesis, it has not been demonstrated to do so.[3] Third, unlike most *N*-methylated amino sugar-containing natural product gene clusters,[16,17] the nogalamycin gene cluster encodes two *N*-methyltransferases, SnogA and SnogX.[1] The function of neither methyltransferase has been demonstrated, and, further, it is not known whether both are required for methylation of nogalamine or, if both are required, what their order of reaction is.[5]

Because anthracyclines with glycosylated D rings are not well-represented amongst the anthracyclines in current clinical use, a better understanding of how nogalamine is biosynthesized and becomes dually linked to ring D of the aglycon could provide novel approaches to overcome the anthracycline class limitations of fatal cardiotoxicity and the development of tumor resistance.[7,18,19] To address some of these issues, we sought to establish the sequence of reactions giving rise to TDP-L-2'-deoxy nogalamine, the likely precursor to the amino sugar found in nogalamycin, as well as explore the reaction sequence leading to the formation of the nogalamine C2'-aglycon C2 bond. Herein we report the *in vitro* reconstitution of the TDP-L-2'-deoxy nogalamine biosynthetic pathway and demonstrate the functions of SnogF, SnogG, SnogA, and SnogX. Furthermore, evidence to support the feasibility of anthracycline biosynthetic engineering in *Streptomyces peucetius* is provided.

## 4.2. EXPERIMENTAL PROCEDURES

### *4.2.1. Chemicals, Reagents, and Materials*

All chemicals and reagents were purchased from Sigma-Aldrich (St. Louis, MO) or Fisher Scientific (Pittsburg, PA) and used without further purification unless otherwise noted. Laboratory consumables such as microcentrifuge tubes, conical tubes, and pipette tips were sourced from Fisher Scientific or VWR International (Radnor, PA).

*E.coli* DH5α was acquired from Bethesda Research Laboratories (Muskegon, WI), pET28b(+) vector DNA was obtained from Novagen (Madison, WI), KOD Hot Start® DNA Polymerase Kit was purchased from EMD Chemicals (Madison, WI), restriction endonucleases and T4 DNA Ligase Kit were obtained from New England Biolabs (Ipswich, MA), and all reagents and apparatus for SDS-PAGE were purchased from Bio-Rad (Hercules, CA). 1 KB Plus DNA ladder, pCRBlunt vector, and *E. coli* BL21 (DE3) were obtained from Invitrogen (Carlsbad, CA). The Ni-NTA agarose, MiniPrep Kits, and QIAquick Gel Extraction Kits were products of QIAgen (Valencia, CA).

### *4.2.2. Preparation of TDP-D-glucose*

TDP-D-glucose (**11**) was enzymatically synthesized from α-D-glucose-1-phosphate (**10**) using a reported method.[20] The reaction was quenched by filtration through an Amicon® stirred cell and 10 kDa molecular weight cut-off membrane from EMD Millipore under $N_2$. The filtrate was lyophilized to dryness, and the remaining residue was dissolved in ~4 mL deionized $H_2O$ (di$H_2O$) and loaded onto a column of P2 resin from Bio-Rad (Hercules, CA) pre-equilibrated with di$H_2O$. The column was eluted with di$H_2O$ using a

peristaltic pump from Pharmacia (Uppsala, Sweden). Fractions strongly absorbing light at 267 nm were further analyzed by high performance liquid chromatography (HPLC) using a Dionex® CarboPac PA1 analytical column (4 x 250 mm) and a CarboPac PA1 guard column (4 x 50 mm) from ThermoFisher Scientific (West Palm Beach, FL) and the detector was set at 267 nm. The binary mobile phase was comprised of di$H_2O$ (eluent A) and 0.5 M aqueous ammonium acetate (eluent B). The elution program was as follows: 5-20%B over 15 min, 20-60%B over 20 min, 60-100%B over 2 min, hold 3 min, 100-5%B over 5 min, using a flow rate of 1 mL/min. Fractions of >95% purity were pooled and lyophilized to yield TDP-glucose.

*4.2.3. Cloning of SnogF, SnogG, SnogA, and SnogX Genes*

*Streptomyces nogalater* NRRL 3035 was obtained from the NRRL Agricultural Research Service Culture Collection (Peoria, IL) and propagated on YM agar (0.3% yeast extract, 0.3% malt extract, 0.5% peptone, 1% dextrose, 2% agar) or in tryptic soya broth (TSB, pre-mixed powder from Sigma-Aldrich) for liquid cultures, and stored as frozen mycelia or spore suspensions in 20% glycerol at -80 °C. *S. nogalater* was grown and maintained according to standard protocols found in *Practical Streptomyces Genetics*.[21] Whole genomic DNA was isolated using the DNeasy Blood and Tissue Kit from QIAgen (Valencia, CA) according to the manufacturer's instructions using the recommended pre-treatment for Gram positive bacteria. All protocols relating manipulation of recombinant DNA were from *Molecular Cloning: A Laboratory Manual*.[22] Oligonucleotides custom-made by Sigma-Aldrich. The genes *snogF*, *snogG*, *snogA*, and *snogX*, were cloned from

*S. nogalater* NRRL 3035 genomic DNA with engineered 5'-*NdeI* and 3'-*NheI* sites (in bold) using oligonucleotides listed in Table 1.

| Gene | Sequence (5'→3') |
|---|---|
| *snogF* | Forward:<br>**CATATG**GAGTCACGCACGCTTCTCGTC<br>Reverse:<br>TT**GCTAGC**TCATCCGGCGTACAGGTCGTG |
| *snogG* | Forward:<br>**CATATG**ACCGCCGGGATACCGCGGAC<br>Reverse:<br>**GCTAGC**TCATAAGGCTGACTCCCGATAGC |
| *snogA* | Forward:<br>**CATATG**GTGTACGGCCGGGAACTGGC<br>Reverse:<br>ATAATAATA**GCTAGC**TTACCCGGCCCCGCGCACCCC |
| *snogX* | Forward:<br>**CATATG**GTGACGCGTGTGTACGGCAC<br>Reverse:<br>AATTAATAATA**GCTAGC**TCACCCCGCCCGCCCGGACTTC |

**Table 4.1:** Sequence of oligonucleotide primers used to clone genes from *S. nogalater*

Polymerase chain reaction (PCR) was carried out in 0.2 mL thin-walled mini tubes using the following reaction composition with KOD DNA Polymerase Kit reagents: 2.5 µL 10x buffer, 2.5 µL dNTPs mix, 2.5 µL DMSO, 2 µL $MgSO_4$ solution (1 mM final concentration, 0.75 µL of each primer in $ddH_2O$ (3 µM final concentration), 1 µL template DNA (~12 ng) in QIAgen buffer AE, and 12.5 µL $ddH_2O$. PCR was carried in an Eppendorf Thermocycler Gradient® as follows. After heating the lid of the thermocycler to 105 °C and an initial 2 min incubation at 95 °C, 31 cycles were carried out consisting of 95 °C for 20 s, lowest primer melting temperature for 10 s, and 70 °C for 15 s/kb target length, and then 70 °C for 10 min. PCR products were purified by agarose gel electrophoresis and then cloned into pCRBlunt for Sanger sequencing at the Core Facilities of the Institute for

Cell and Molecular Biology at the University of Texas at Austin, Austin, TX. Correctly sequenced genes were excised from pCRBlunt, and the gel-purified DNA was ligated between their respective sites in pET28b(+) for expression with N-terminal $His_6$-tags.

*4.2.4. Expression and Purification of SnogF, SnogG, SnogA, SnogX, RfbB, TylX3, TylC2, and EvaB*

The pET28(+) constructs were used to transform *E. coli* BL21 (DE3) for expression. The recombinant strains were grown in lysogeny broth (LB) supplemented with 30 μg/mL kanamycin. After growth at 37 °C until $OD_{600}$ ~0.6, cultures harboring the *snogA*:pET28(+) and *snogF*:pET28b(+) constructs were induced with 100 μM IPTG and allowed to grow at 16 °C for 16 h; the culture harboring the *snogX*:pET28b(+) construct was induced with 200 μM IPTG and allowed to grow at 30 °C for 16 h; and the culture harboring the *snogG*:pET28b(+) construct was induced with 200 μM IPTG and allowed to grow at 25 °C for 16 h. The cultures were harvested by centrifugation at 4500 x *g* at 4 °C, and the cell pellets (typically ~5 g/L) were collected and stored at -80 °C until purification.

For purification, the cell pellets were thawed in room temperature water and resuspended in two volumes of either lysis buffer A, composed of 50 mM $NaH_2PO_4$, 300 mM NaCl, 10 mM imidazole, and 20% glycerol, pH 8.0, (SnogA, SnogF, and SnogX) or lysis buffer B, composed of 50 mM $NaH_2PO_4$, 300 mM NaCl, 10 mM imidazole, 20% glycerol, 10 mM L-Arg, 10 mM L-Glu, and 0.5 mM DTT; pH 8.0 (SnogG). Egg white lysozyme was added to a concentration of 1 mg/mL, and the cell suspensions were allowed to sit on ice for 30 min with intermittent stirring. Sonication was performed at maximum power with 30 s bursts and one min cooling periods between bursts. The sonicated suspensions were then centrifuged at 20000 x *g* at 4 °C for 40 min, and the supernatants

were applied to pre-packed, lysis buffer-equilibrated columns of 6.5 mL of Ni-NTA resin and allowed to flow through. The columns were then washed with ~100 mL of either wash buffer A, which has the same composition of lysis buffer A but has 20 mM imidazole, or wash buffer B, which has the same composition of lysis buffer B but has 20 mM imidazole. Finally, the proteins were eluted with either elution buffer A, which has the same composition of lysis buffer A but has 250 mM imidazole, or elution buffer B, which has the same composition as lysis buffer B but has 250 mM imidazole. After dialysis into dialysis buffer A, which is composed of lysis buffer A but with no imidazole, or dialysis buffer B, which is composed of lysis buffer B but with no imidazole, the pooled column eluates were concentrated to 5-8 mL using an Amicon® stirred cell and 10 kDa molecular weight cut-off membrane from EMD Millipore (St. Louis, MO) under $N_2$. The proteins were stored as 1 mL aliquots at -80 °C.

RfbB, a 4',6'-dehydratase cloned from *Salmonella typhimurium*;[23] TylX3, a 2',3'-dehydratase cloned from *Streptomyces fradiae*;[24] TylC2, a 4'-keto reductase cloned from *S. fradiae*;[25] and EvaB, a 4'-aminotransferase cloned from *Amycolatopsis orientalis*[26] were expressed and purified as previously reported.

*4.2.5. Enzymatic Assays and HPLC Analyses Thereof*

Assays of the functions of SnogF, SnogG, SnogA, and SnogX were carried out in 50 mM Tris HCl buffer (pH 8.0) and set up in two steps (Figure 4.4). The first step was used to generate TDP-4-keto-6-deoxy glucose (**12**) and was composed of 5 mM TDP-glucose and 15 μM RfbB. After incubation at 37 °C for one hour, the nascent product **12**

was subjected to the second step of the assay was used to generate **14**-**16**. The reaction mixture from the first step was made 1.5 mM in PLP, 62.5 mM in L-Glu, 5 mM in NADPH, 20 μM in EvaB, 20 μM in SnogF, 30 μM in SnogG, and the addition of TylX3 to 10 μM initiated the second step of the assay. After incubation at 30 °C for one hour, samples were drawn, and after filtration through Amicon® YM-10 centrifugal filters from EMD Millipore, the filtrates were analyzed at 267 nm by analytical-scale HPLC using the same column and elution program described earlier to check the purity of P2 column fractions. Under these conditions, **14** elutes at ~11 min, **15** elutes at ~12 min, and **16** elutes at ~9 min.

To demonstrate the observed shunt reaction, the second step of the reaction was made 5 mM in NADPH, 20 μM in SnogF and 30 μM in SnogG or TylC2. To analyze the sequence of methylations by SnogA and SnogX, one or both of these enzymes was added to the second step of the reaction at a concentration of 20 μM along with *S*-adenosylmethionine (SAM) at a concentration of 5 mM to generate compounds **17** and **18**. Using the same HPLC conditions as before, **17** and **18** eluted together at ~7.8 min. Resolution of **17** and **18** was achieved by using the same mobile phase as before but beginning the HPLC elution program with a gradient from 0% to 30% eluent B.

To obtain samples for mass spectrometry, eluted peaks of interest were collected from repeated injections under analytical conditions, pooled and lyophilized, and the resulting residues were dissolved in water. Mass spectrometry was performed at the Mass Spectrometry and Proteomics Facility of the Department of Chemistry and Biochemistry at the University of Texas at Austin, Austin, TX.

HR-MS ($ESI^-$) *m/z* for (**16**) calc.'d $C_{16}H_{26}N_3O_{13}P_2^-$ $[M-H]^-$, 530.0946, found 530.0950

HR-MS ($ESI^-$) *m/z* for (**17**) calc.'d $C_{17}H_{28}N_3O_{13}P_2^-$ $[M-H]^-$, 544.1103, found 544.1087

HR-MS ($ESI^-$) *m/z* for (**18**) calc.'d $C_{18}H_{30}N_3O_{13}P_2^-$ $[M-H]^-$, 558.1259, found 558.1249

Samples for $^1$H-NMR were collected from preparative scale (10 mL) reactions performed using the same conditions and concentrations of reactants as were used for the analytical-scale reactions. After completion, the reaction was filtered using an Amicon® stirred cell and 10 kDa molecular weight cut-off membrane from EMD Millipore. The filtrates were purified using HPLC with a Dionex® CarboPac PA1 semi-preparative column (9 x 250 mm) and semi-preparative CarboPac PA1 guard column (9 x 50 mm) from ThermoFisher Scientific. The binary mobile phase and elution program used were the same as were used for analytical-scale reactions, except a flow rate of 3 mL/min was used. Pooled eluted peaks were lyophilized, and the residues were dissolved in $D_2O$. $^1$H-NMR data were obtained using a Varian® Inova 500 MHz instrument at the Nuclear Magnetic Resonance Facility of the Department of Chemistry and Biochemistry at the University of Texas at Austin, Austin, TX.

Compound **16**: $^1$H NMR ($D_2O$, 500 MHz) δ 7.61 (s, 1H, 6-H), 6.21 (m, 1H, 1"-H), 5.18 (dd, *J*=5, *J*=7.5, 1H, 1'-H), 4.46 (m, 1H, 3"-H), 4.05 (m, 3H, 4"-H, 5"-H), 3.69 (m, 1H, 5'-H), 3.44 (m, 2H, 3'-H, 4'-H), 2.22 (m, 2H, 2"-H), 2.10 (m, 1H, 2'-H), 1.81 (s, 3H, 5-Me), 1.68 (m, 1H, 2'-H), 1.17 (d, *J*=6.5, 3H, 5'-Me)

Compound **20**: $^1$H NMR ($D_2O$, 500 MHz) δ 7.62 (s, 1H, 6-H), 6.22 (t, *J*=5.5, *J*=6, 1H, 1"-H), 5.34 (dd, *J*=20.5, *J*=1.5, 1H, 1'-H), 4.43 (m, 1H, 3"-H), 4.39 (m, 1H, 2'-H), 4.03 (m, 1H, 4"-H), 3.88 (m, 2H, 5"-H), 3.68 (m, 1H, 5'-H), 3.56 (m, 1H, 3'-H), 3.52 (m, 1H, 4'-H), 2.2-2.3 (m, 2H, 2"-H), 1.18 (d, *J*=5.5, 3H, 5'-Me)

*4.2.6. Creation of Plasmids pL234 and pL2W34-5 and Transformation of Streptomyces peucetius ATCC 29050 and Streptomyces peucetius WHM1628*

*Streptomyces peucetius* ATCC 29050 and *Streptomyces peucetius* WHM1628, a *dnrQ⁻/dnrS⁻* mutant of *S. peucetius* ATCC 29050, were kind gifts from the laboratory of the late Dr. Richard Hutchinson. Unless stated otherwise laboratory methods pertaining to *Streptomyces* as well as growth media are described in *Practical Streptomyces Genetics.*[21] These strains were propagated on R2 or R5 agar plates or in R5 liquid (prepared as described for R5 solid medium by Kieser, et al.[21] without agar), and stored as frozen mycelia suspensions in 20% glycerol at -80 °C. Oligonucleotides were obtained from Sigma-Aldrich. The genes *snoaL2* and *snoaW* were cloned from *S. nogalater* genomic DNA with engineered ribosomal binding sites (RBS) based on a consensus *Streptomyces* RBS sequence[21] using the primers shown in Table 2. The engineered restriction sites are bolded. The PCR reactions were set up and performed as described earlier.

| Gene | Sequence (5'→3') |
|---|---|
| *snoaL2* | Forward: *XbaI*<br>**TCTAGA**GGAGGTGCCGCCATGTCAACGACCGCGAAC<br>Reverse: *HindIII*<br>**AAGCTT**TCACAGTCCGTCCGGCACCACGC |
| *snoaW* | Forward: *BamHI*<br>**GGATCC**AGGAGGTGCCGCCATGACAGTATTGGTAAC<br>Reverse: *XbaI*<br>**TCTAGA**TCACCGGAAGTCCGCTCGGTGGTC |

**Table 4.2:** Sequences of oligonucleotide primers used to clone *snoaL2* and *snoaW* for construction of plasmids pL2345 and pL2W34-5. RBS and spacer are underlined.

The *snoaL2* and *snoaW* PCR products were purified by agarose gel electrophoresis, and after gel extraction, each was cloned into pCRBlunt. The pCRBlunt constructs were Sanger sequenced at the Core Facilities of the Institute for Cell and Molecular Biology at the University of Texas at Austin, Austin, TX. The correctly sequenced genes were excised with the restriction endonucleases indicated in Table 2 and gel purified.

For the construction of plasmid pL234, digested *snoaL2* was cloned into the respective sites of the *Streptomyces*/*E. coli* shuttle vector pSE34.[23] For the construction of plasmid pL2W34-5, plasmid pL234 was digested with *BamHI* and *XbaI*, and the digested *snoaW* was cloned into these sites.

*4.2.7. Culture Conditions for Anthracyclines Production in S. peucetius ATCC 29050 and S. peucetius WHM1628*

10 mL cultures of *S. peucetius* ATCC 29050 and *S. peucetius* WHM1628 inoculated with 0.2 mL frozen mycelial suspension were grown in capped culture tubes (25 x 150 mm) with glass beads at 30 °C with rotary shaking at 200 RPM for 2-3 days. 1 mL of each culture was used to inoculate a respective culture of 100 mL R5 liquid medium supplemented with 0.5% glycine in 1 L baffled flasks. These cultures were grown at 30 °C with rotary shaking at 200 RPM for 2-3 days, and then protoplasts of *S. peucetius* ATCC 29050 and *S. peucetius* WHM1628 were prepared. Protoplasts were transformed with 3-5 µg of plasmid DNA using 25% PEG 1000 in P-buffer and plated on R2 agar. After incubating the plates overnight at 30 °C, they were overlaid with 1.25 µg each of thiostrepton for selection, using a 50 mg/mL stock solution in DMSO. The antibiotic-

treated plates were incubated at 30 °C for 4-5 more days, or until the bright yellow transformants appeared. Pieces of the R2 plates containing *S. peucetius* ATCC 29050 or *S. peucetius* WHM1628 transformants were cut using a flamed inoculating loop and aseptically transferred to culture tubes containing 10 mL liquid R5 medium with thiostrepton (50 µg/mL) and glass beads. These tubes were then capped and grown at 30 °C with rotary shaking at 200 RPM. After 3-4 days, 1 mL of these cultures was used to inoculate 200 mL of a slight variation of the production medium described by Dekleva, et al.[24] containing ($L^{-1}$) 30 g dextrose, 10 g Traders® Protein from ADM (Chicago, IL), 3 g NaCl, 3 g $CaCO_3$, 10 mL of the trace elements solution described in *Practical Streptomyces Genetics*[21] and thiostrepton (12.5 µg/mL). These cultures were grown at 30 °C with rotary shaking at 220 RPM for 10-12 days.

### *4.2.8. Isolation of Anthracyclines from S. peucetius ATCC 29050 and S. peucetius WHM1628*

To isolate anthracyclines, the cultures were centrifuged at 5000 x *g* at 4 °C for 30 min. The supernatant was removed, adjusted to pH 2.0 using phosphoric acid, and allowed to stir with 40 g Diaion™ HP-20 resin (Mitsubishi Chemical; White Plains, NY) at room temperature for 2 h. The cell pellet was collected and allowed to stir with 500 mL methanol at room temperature for 2 h as well. After 2 h, the Diaion™ resin was collected by filtration through a Büchner funnel, transferred to a flask and allowed to stir with 500 mL fresh methanol overnight at 4 °C; the cell pellet slurry was also filtered, and the filtrate was combined with the first resin elution and stored at 4 °C. The next day, the resin was

collected again, transferred back to a flask and eluted once more with 500 mL fresh methanol by stirring for an additional 2 h at 4 °C. The resin was then removed, and the filtrate was combined with the earlier elution fractions. The total eluate (~2 L) was then concentrated to dryness through rotary evaportation.

## 4.3. RESULTS AND DISCUSSION

### *4.3.1. Biosynthetic Study of TDP-nogalamine*

Based on our analysis of the nogalamycin biosynthetic gene cluster, comparison with canonical deoxy sugar pathways, and results from our assays, we constructed a putative pathway for the biosynthesis of TDP-L-2'-deoxy nogalamine (**18**, Scheme 1). RfbB and TylX3, a 4',6'-dehydratase and a 2',3'-dehydratase, respectively, were used in the assays in place of the equivalent enzymes from the nogalamycin biosynthetic gene cluster due to their ease of expression and purification,[20] and their placements in 2'-deoxy sugar reaction pathways have been thoroughly investigated.[14]

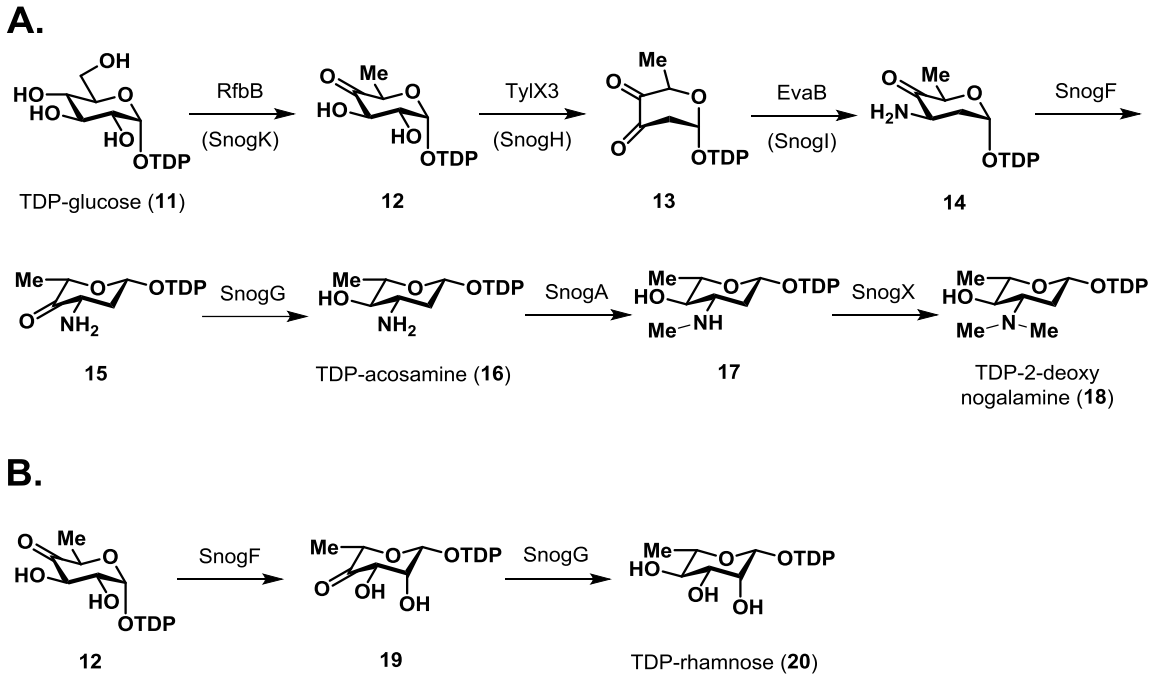



**Figure 4.4: A.** Proposed biosynthetic pathway of TDP-2-deoxy nogalamine; **B.** Observed shunt reaction

The first point of uncertainty in our putative pathway for the biosynthesis of TDP-2'-deoxy nogalamine is the stereochemical outcome of the predicted 3'-aminotransferase, SnogI. Structures featuring axially[5] and equatorially[1] placed amino groups have been proposed in the literature, but based on the structure of nogalamycin and by analogy to other deoxy amino sugar pathways,[25,26] we expected the amino group to be installed equatorially. Due to difficulties in expressing *snogI* and purifying its product, we utilized EvaB from the TDP-L-*epi*-vancosamine pathway, which is known to produce an equatorially placed amino group.[25] As mentioned before, the nogalamine biosynthetic gene cluster encoded two deoxy sugar C4'-keto reductases, SnogC and SnogG. To provide

insight into the most likely enzyme involved in biosynthesis of the amino sugar, we performed phylogenetic analysis of SnogC and SnogG. Because SnogG was found to clade with other amino sugar C4'-keto reductases, this enzyme was chosen as the putative TDP-L-nogalamine 4'-keto reductase instead of SnogC. The assays performed and subsequent analyses supported these assertions.

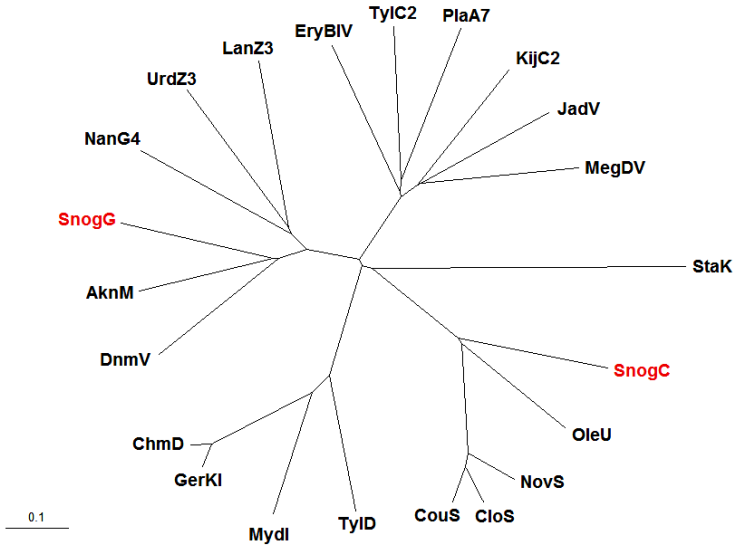



**Figure 4.5:** Phylogenetic relationship of SnogG and SnogC to other known deoxy sugar C4'-keto reductases. AknM: aclacinomycin gene cluster from *Streptomyces galilaeus* ATCC 31615; DnmV: daunomycin gene cluster from *Streptomyces peucetius* ATCC 29050. Figure created with TreeView, version 1.6.6, available from *http://taxonomy.zoology.gla.ac.uk/rod/treeview.html*

When **12** was reacted with TylX3 and EvaB, a new compound, **14**, was formed. The addition of SnogG and SnogF resulted in formation of 16. Surprisingly, it was found that SnogF and SnogG also accept **12** as a substrate to yield **20**, TDP-rhamnose. This is in

accord with the proposal that SnogF acts in both the nogalamine and nogalose pathways and indicates that SnogF and SnogG could be substrate-flexible. Interestingly, SnogC, which we propose to be involved in nogalose biosynthesis, has been used in an *in vivo* approach to generate deoxy amino sugars.[27] To more closely observe this ostensible shunt pathway, **12** was reacted with SnogG and SnogF or with TylC2 and SnogF. TylC2 is a 4'-keto reductase from the TDP-L-mycarose pathway previously purified by our laboratory[28,29] that is expected to generate an equatorially placed alcohol, as predicted for SnogG. The formation of 20 was only seen when 12 was reacted with SnogG and SnogF.

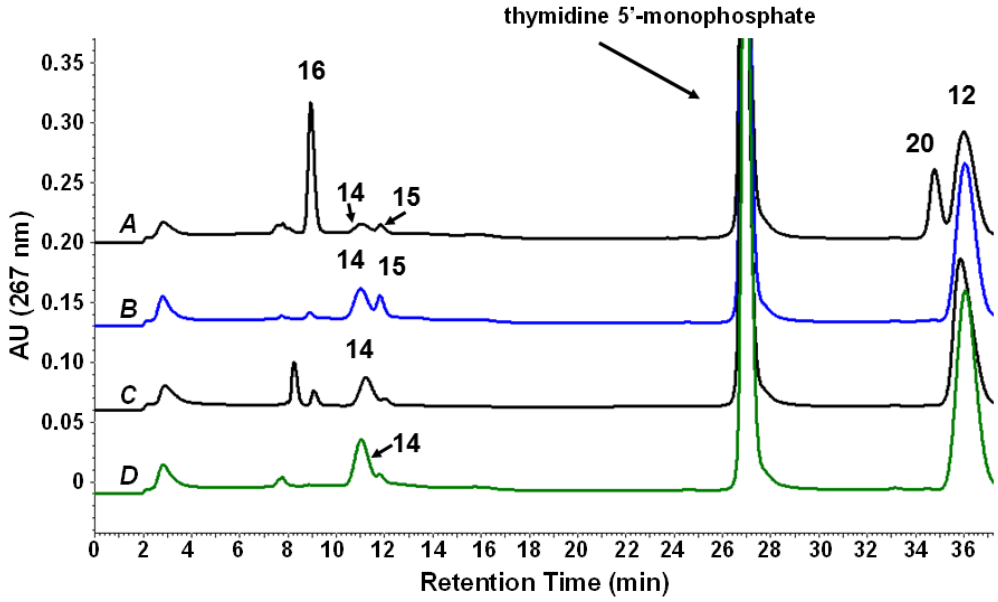



**Figure 4.6:** Validity of proposed pathway. Trace A: RfbB, TylX3, EvaB, SnogF, and SnogG; Trace B: RfbB, TylX3, EvaB, SnogF; Trace C: RfbB, TylX3, EvaB, SnogG; Trace D: RfbB,TylX3, EvaB

Mass spectrometry data support the assignments of peak for **16**. The $^{1}$H-NMR data collected for **16** are in good agreement with those previously published for TDP-L-

acosamine.[30] These data support the roles we predict for SnogF and SnogG—a 3',5'-epimerase and a 4'-keto reductase capable of accepting an amino sugar. Furthermore, based on our results, we expect that SnogI results in an equatorially placed amino group. $^{1}$H-NMR data for **20** agree with those published for TDP-L-rhamnose,[31-33] supporting our predicted role for SnogF but contrary to a previous prediction that SnogF is only a 5'-epimerase,[3] a function seemingly necessary for TDP-L-nogalose biosynthesis. However, it can be appreciated that SnogF is only *effectively* a 5'-epimerase in the biosynthesis of nogalose because there is no 3' proton available for abstraction since this position is doubly substituted.[33,34] This observation is in agreement with the apparent substrate flexibility of SnogF.

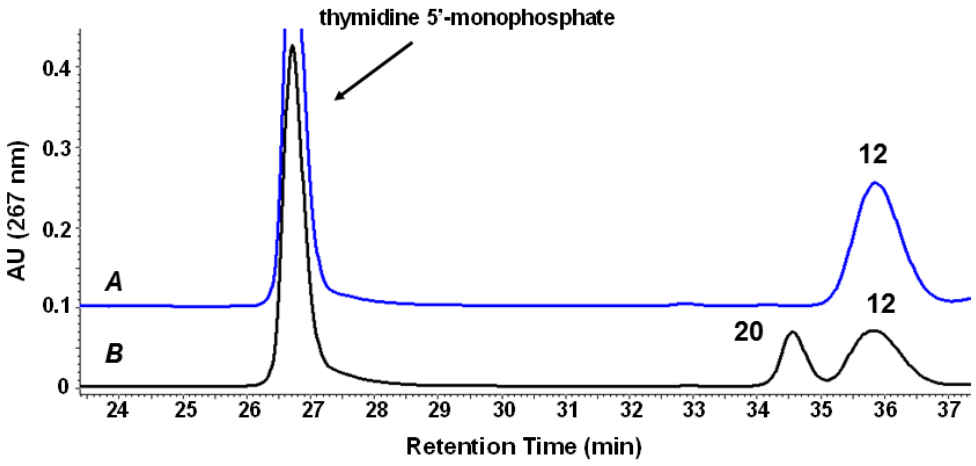



**Figure 4.7:** Promiscuity of SnogG; Trace A: RfbB, SnogF, TylC2; Trace B: RfbB, SnogF, SnogG

Based on the available literature pertaining to nogalamycin, methylation is expected to be a post-glycosylation modification requiring two separate *N*-methyltransferases, SnogA and SnogX,[5] but this is in contrast to the established biosyntheses of other *N*-dimethylated deoxy amino sugars, where methylation occurs prior to glycosylation, and dimethylation is accomplished through the iterative action of a single methyltransferase.[16,25,26]

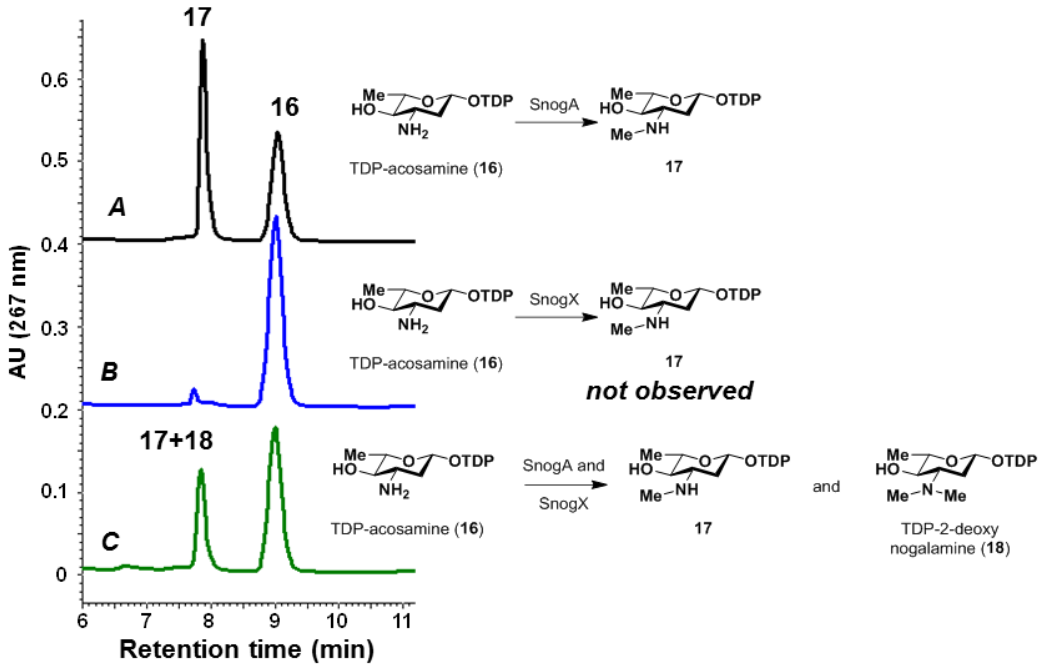



**Figure 4.8:** *N*-methyltransferase Assays; Trace A: RfbB, TylX3, EvaB, SnogF, SnogG, SnogA; Trace B: RfbB, TylX3, EvaB, SnogF, SnogG, SnogX; Trace C: RfbB, TylX3,EvaB, SnogF, SnogG, SnogA, SnogX

Therefore, we predicted that the methylations actually take place at the TDP-deoxy sugar stage, but the roles of the methyltransferases and the sequence of methylation are not clear. When SnogA was the only methyltransferase present in the second stage of the assays, monomethylated compound **17** resulted. However, when SnogA was replaced with

SnogX, only scant product formation was seen. The addition of SnogA and SnogX together in assays results in a mixture of **17** and the dimethylated amino sugar **18**, TDP-L-2'-deoxy nogalamine.

To more clearly characterize the sequence of activity of SnogA and SnogX, assays were performed with SnogA alone, resulting solely in **17**, and addition of SnogX after a one hour incubation resulted in the formation of **17**. These assignments were supported with mass spectrometry data. Thus, SnogA and SnogX are required to generate dimethylated amino sugar **18**, TDP-L-2'-deoxy nogalamine, via two sequential monomethylations. Furthermore, SnogA performs the first methylation, and SnogX performs the second methylation.

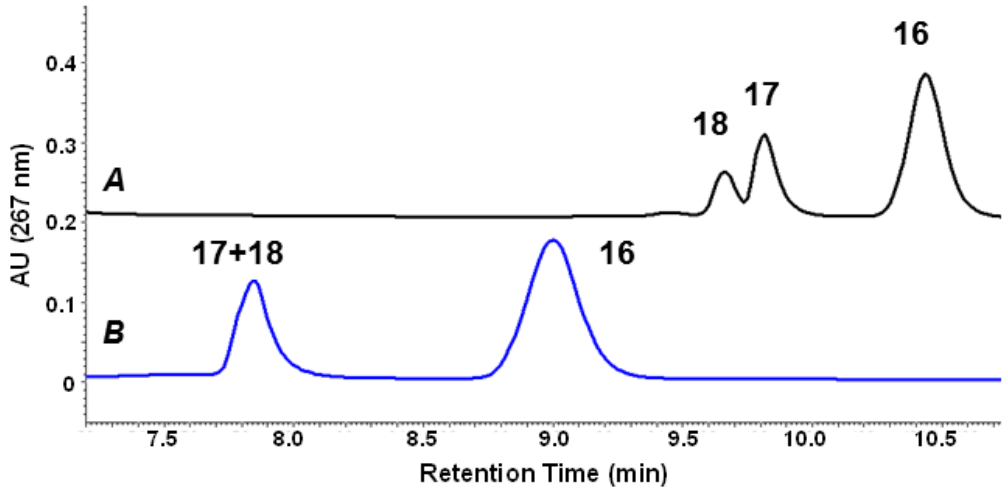



**Figure 4.9:** Resolution of *N*-methylated amino sugars; Trace A: 0.5 M $NH_4OAc$ gradient from 0% to 30% over 15 min; Trace B: 0.5 M $NH_4OAc$ gradient from 5% to 20% over 15 min

*4.3.2. Anthracycline Bioengineering in Streptomyces peucetius*

Studies on the biosynthesis of the E-F ring system of nogalamycin or other extended ring systems are complicated by the need for C1-hydroxylated anthracycline aglycons.[3,4,35] Due to the preponderance of reactive positions on an anthracycline aglycon, a semisynthetic approach to the installation of a hydroxyl at C1 using a natural product anthracycline is not feasible.[36] Thus, creating a C1-hydroxylated anthracycline through chemical means would involve total synthesis. Although total synthesis has been used to generate many anthracycline analogs, the reaction schemes are lengthy and time consuming.[37] Alternatively, bioengineering of natural anthracycline production could take advantage of the producing organisms' abilities to generate complex, synthetically inaccessible structures.[3,4,35] *S. nogalater* has been reported to be recalcitrant to genetic manipulation,[38] but since the enzymes responsible for introducing the C1-hydroxyl have been identified (SnoaL2 and SnoaW), our approach focused on the introduction of the requisite genes into another genetically tractable host. Through the course of extensive study regarding the biosyntheses of daunorubicin and doxorubicin, *S. peucetius* proved to be amenable to transformation.[39-42]

Nogalamycin is distinguished from daunorubicin and doxorubicin by the presence of a methyl ester at the C10 position, as this substituent is removed in later biosynthetic steps of the latter two compounds.[39-42] However, ε-rhodomycinone, an early precursor to the daunorubicin aglycon, maintains the C10 methyl ester, so for our initial attempts to bioengineer the production of C1-hydroxylated anthracyclines, *S. peucetius* WHM1628, a

blocked mutant which accumulates ε-rhodomycinone (**21**),[40] was used as a host. First, *snoaL2* was cloned with an artificial *Streptomyces* ribosomal binding site (RBS) downstream from the *ermE** promoter of pSE34 to yield plasmid pL234. However, introduction of this plasmid into *S. peucetius* WHM1628 failed to provide the desired hydroxylated compound. Unexpectedly, during this attempt, Siitonen, et al. showed that SnoaL2 requires the quinone reductase SnoaW to perform hydroxylation.[43] Thus, *snoaW* was cloned with the same artificial RBS used for pL234 upstream from *snoaL2* to yield plasmid pL2W34-5.

Unfortunately, introduction of pL2W34-5 into *S. peucetius* WHM1628 did not provide C1-hydroxylated ε-rhodomycinone either. Once again, in the midst of this work, Siitonen, et al. demonstrated *in vivo* hydroxylation of nogalamycinone, the nogalamycin aglycon, in a heterologous host required prior glycosylation of the C7-hydroxyl.[5] With this in mind, pL2W34-5 was introduced into *S. peucetius* ATCC29050 strain.

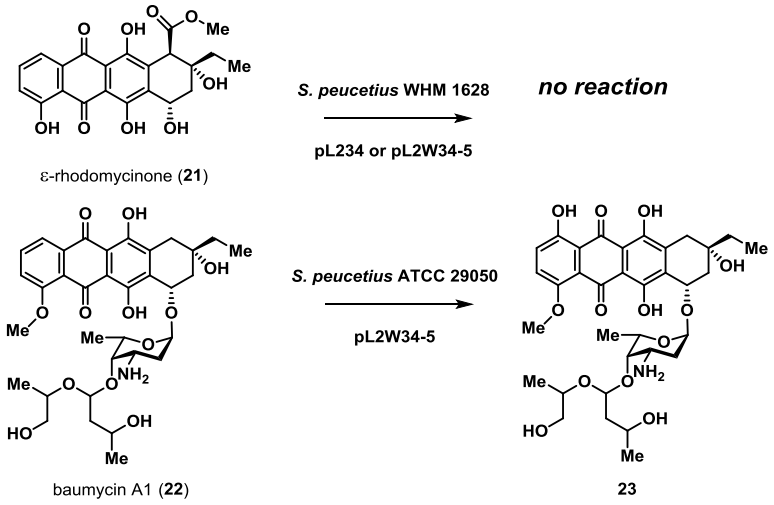



**Figure 4.10:** C1 hydroxylation of baumycin A1; A. No reaction with ε-rhodomycinone; B. hydroxylated baumycin A1

This "wild-type" strain is renowned for its ability to produce daunorubicin, but its major product is actually a higher glycoside of daunorubicin, baumycin A1, which is hydrolyzed to increase the yield of daunorubicin.[44] Gratifyingly, mass spectrometric analysis of a culture extract of the transformed *S. peucetius* ATCC29050 strain revealed the presence of C1-hydroxy-baumycin A1 (**23**). This compound was not detected in a control culture of untransformed *S. peucetius* ATCC29050. Thus, pL2W34-5 could be useful in procuring C1-hydroxylated anthracyclines.

## 4.4. CONCLUSIONS

Our work has provided the first insights into TDP-nogalamine biosynthesis. We have demonstrated the activities of SnogF, SnogG, SnogA, and SnogX, thereby laying the groundwork for future investigations into the biosynthesis of the deoxy sugars decorating nogalamycin. Importantly, we have shown that dimethylation of TDP-nogalamine does indeed require both putative *N*-methyltransferases, SnogA and SnogX, encoded in the nogalamycin biosynthetic gene cluster, we have established the sequence of methylations performed by SnogA and SnogX, and we have demonstrated that dimethylation can occur at the TDP-deoxy sugar stage. Our results also support the prediction that SnogF is shared between the biosynthetic pathways for TDP-nogalamine and TDP-nogalose. Taken together, there is now some evidence that TDP-deoxy sugar biosynthesis in the nogalamycin biosynthetic pathway may be amenable to manipulation. Furthermore, initial bioengineering efforts in *S. peucetius* ATCC29050 have demonstrated the feasibility of using this strain to produce C1-hydroxylated anthracyclines. In all, the groundwork has

been laid for further study into the biosynthesis of anthracyclines with extended ring systems. It is hoped future developments in this area could provide strategies to circumvent anthracycline toxicity and the development of tumor resistance.

## 4.6. REFERENCES


(1) Torkkell, S.; Kunnari, T.; Palmu, K.; Mantsala, P.; Hakala, J.; Ylihonko, K. The entire nogalamycin biosynthetic gene cluster of Streptomyces nogalater: characterization of a 20-kb DNA region and generation of hybrid structures. *Mol Genet Genomics.* **2001**, *266*, 276.

(2) Torkkell, S.; Ylihonko, K.; Hakala, J.; Skurnik, M.; Mantsala, P. Characterization of Streptomyces nogalater genes encoding enzymes involved in glycosylation steps in nogalamycin biosynthesis. *Mol Gen Genet.* **1997**, *256*, 203.

(3) Metsa-Ketela, M.; Niemi, J.; Mantsala, P.; Schneider, G. Anthracycline Biosynthesis: Genes, Enzymes and Mechanisms. *Topics in Current Chemistry.* **2008**, *282*, 101.

(4) Niemi, J.; Metsa-Ketela, M.; Schneider, G.; Mantsala, P. Biosynthetic Anthracycline Variants. *Topics in Current Chemistry.* **2008**, *282*, 75.

(5) Siitonen, V.; Claesson, M.; Patrikainen, P.; Aromaa, M.; Mantsala, P.; Schneider, G.; Metsa-Ketela, M. Identification of late-stage glycosylation steps in the biosynthetic pathway of the anthracycline nogalamycin. *Chembiochem.* **2011**, *13*, 120.

(6) Williams, L. D.; Egli, M.; Qi, G.; Bash, P.; van der Marel, G. A.; van Boom, J. H.; Rich, A.; Frederick, C. A. Structure of nogalamycin bound to a DNA hexamer. *Proc Natl Acad Sci U S A.* **1990**, *87*, 2225.

(7) Asche, C. Antitumour quinones. *Mini Rev Med Chem.* **2005**, *5*, 449.

(8) Liaw, Y. C.; Gao, Y. G.; Robinson, H.; van der Marel, G. A.; van Boom, J. H.; Wang, A. H. Antitumor drug nogalamycin binds DNA in both grooves simultaneously: molecular structure of nogalamycin-DNA complex. *Biochemistry.* **1989**, *28*, 9913.

(9) Bhuyan, B. K.; Blowers, C. L.; Shugars, K. D. Lethality of nogalamycin, nogalamycin analogs, and adriamycin to cells in different cell cycle phases. *Cancer Res.* **1980**, *40*, 3437.

(10) Brown, T. D.; Donehower, R. C.; Grochow, L. B.; Rice, A. P.; Ettinger, D. S. A phase I study of menogaril in patients with advanced cancer. *J Clin Oncol.* **1987**, *5*, 92.

(11) Sessa, C.; Gundersen, S.; ten Bokkel Huinink, W.; Renard, J.; Cavalli, F. Phase II study of intravenous menogaril in patients with advanced breast cancer. *J Natl Cancer Inst.* **1988**, *80*, 1066.

(12) Peng, R.; VanNieuwenhze, M. S. A model study for constructing the DEF-benzoxocin ring system of menogaril and nogalamycin via a reductive Heck cyclization. *Org Lett. 14*, 1962.

(13) Hauser, F. M.; Ganguly, D. Synthetic studies on nogarol anthracyclines. Enantioselective total synthesis of an aminohydroxy epoxybenzoxocin. *J Org Chem.* **2000**, *65*, 1842.

(14) Thibodeaux, C. J.; Melancon, C. E., 3rd; Liu, H. W. Natural-product sugar biosynthesis and enzymatic glycodiversification. *Angew Chem Int Ed Engl.* **2008**, *47*, 9814.

(15) Tello, M.; Rejzek, M.; Wilkinson, B.; Lawson, D. M.; Field, R. A. Tyl1a, a TDP-6-deoxy-D-xylo-4-hexulose 3,4-isomerase from Streptomyces fradiae: structure prediction, mutagenesis and solvent isotope incorporation experiments to investigate reaction mechanism. *Chembiochem.* **2008**, *9*, 1295.

(16) Chang, C.; Zhao, L.; Yamase, H.; Liu, H. DesVI: A new member of the sugar N,N-dimethyltransferase family involved in the biosynthesis of desosamine. *Angew Chem Int Ed Engl.* **2000**, *39*, 2160.

(17) Chen, H.; Yamase, H.; Murakami, K.; Chang, C. W.; Zhao, L.; Zhao, Z.; Liu, H. W. Expression, purification, and characterization of two N,N-dimethyltransferases, tylM1 and desVI, involved in the biosynthesis of mycaminose and desosamine. *Biochemistry.* **2002**, *41*, 9165.

(18) Jannazzo, A.; Hoffman, J.; Lutz, M. Monitoring of anthracycline-induced cardiotoxicity. *Ann Pharmacother.* **2008**, *42*, 99.

(19) Barrett-Lee, P. J.; Dixon, J. M.; Farrell, C.; Jones, A.; Leonard, R.; Murray, N.; Palmieri, C.; Plummer, C. J.; Stanley, A.; Verrill, M. W. Expert opinion on the use of anthracyclines in patients with advanced breast cancer at cardiac risk. *Ann Oncol.* **2009**, *20*, 816.

(20) White-Phillip, J.; Thibodeaux, C. J.; Liu, H. W. Enzymatic synthesis of TDP-deoxysugars. *Methods Enzymol.* **2009**, *459*, 521.

(21) Kieser, T., Bibb, M., Buttner, M., Chater, K., Hopwood, D. *Practical Streptomyces Genetics*; The John Innes Foundation: Norwich, 2000.

(22) Sambrook, J., Russell, D. *Molecular Cloning: A Laboratory Manual*; Third

ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, 2001.

(23) Yoon, Y. J.; Beck, B. J.; Kim, B. S.; Kang, H. Y.; Reynolds, K. A.; Sherman, D. H. Generation of multiple bioactive macrolides by hybrid modular polyketide synthases in Streptomyces venezuelae. *Chem Biol.* **2002**, *9*, 203.

(24) Dekleva, M. L.; Titus, J. A.; Strohl, W. R. Nutrient effects on anthracycline production by Streptomyces peucetius in a defined medium. *Can J Microbiol.* **1985**, *31*, 287.

(25) Chen, H.; Thomas, M. G.; Hubbard, B. K.; Losey, H. C.; Walsh, C. T.; Burkart, M. D. Deoxysugars in glycopeptide antibiotics: enzymatic synthesis of TDP-L-epivancosamine in chloroeremomycin biosynthesis. *Proc Natl Acad Sci U S A.* **2000**, *97*, 11942.

(26) Hong, L.; Zhao, Z.; Melancon, C. E., 3rd; Zhang, H.; Liu, H. W. In vitro characterization of the enzymes involved in TDP-D-forosamine biosynthesis in the spinosyn pathway of Saccharopolyspora spinosa. *J Am Chem Soc.* **2008**, *130*, 4954.

(27) Han, A. R.; Park, J. W.; Lee, M. K.; Ban, Y. H.; Yoo, Y. J.; Kim, E. J.; Kim, E.; Kim, B. G.; Sohng, J. K.; Yoon, Y. J. Development of a Streptomyces venezuelae-based combinatorial biosynthetic system for the production of glycosylated derivatives of doxorubicin and its biosynthetic intermediates. *Appl Environ Microbiol.* **2011**, *77*, 4912.

(28) Takahashi, H.; Liu, Y. N.; Chen, H.; Liu, H. W. Biosynthesis of TDP-l-mycarose: the specificity of a single enzyme governs the outcome of the pathway. *J Am Chem Soc.* **2005**, *127*, 9340.

(29) Takahashi, H.; Liu, Y. N.; Liu, H. W. A two-stage one-pot enzymatic synthesis of TDP-L-mycarose from thymidine and glucose-1-phosphate. *J Am Chem Soc.* **2006**, *128*, 1432.

(30) Oberthur, M.; Leimkuhler, C.; Kahne, D. A practical method for the stereoselective generation of beta-2-deoxy glycosyl phosphates. *Org Lett.* **2004**, *6*, 2873.

(31) Graninger, M.; Nidetzky, B.; Heinrichs, D. E.; Whitfield, C.; Messner, P. Characterization of dTDP-4-dehydrorhamnose 3,5-epimerase and dTDP-4-dehydrorhamnose reductase, required for dTDP-L-rhamnose biosynthesis in Salmonella enterica serovar Typhimurium LT2. *J Biol Chem.* **1999**, *274*, 25069.

(32) Madduri, K.; Waldron, C.; Merlo, D. J. Rhamnose biosynthesis pathway supplies precursors for primary and secondary metabolism in Saccharopolyspora spinosa. *J Bacteriol.* **2001**, *183*, 5632.

(33) Dong, C.; Major, L. L.; Srikannathasan, V.; Errey, J. C.; Giraud, M. F.; Lam, J. S.; Graninger, M.; Messner, P.; McNeil, M. R.; Field, R. A.; Whitfield, C.; Naismith, J. H. RmlC, a C3' and C5' carbohydrate epimerase, appears to operate

via an intermediate with an unusual twist boat conformation. *J Mol Biol.* **2007**, *365*, 146.

(34) Dong, C.; Major, L. L.; Allen, A.; Blankenfeldt, W.; Maskell, D.; Naismith, J. H. High-resolution structures of RmlC from Streptococcus suis in complex with substrate analogs locate the active site of this class of enzyme. *Structure.* **2003**, *11*, 715.

(35) Laatsch, H.; Fotso, S. Naturally Occurring Anthracyclines. *Topics in Current Chemistry.* **2008**, *282*, 3.

(36) Silverman, R. *The Organic Chemistry of Enzyme-Catalyzed Reactions*; Revised ed.; Elsevier: Boston, 2002.

(37) Thomas, G. J. In *Recent Progress in the Chemical Synthesis of Antibiotics*; First ed.; Lukacs, G., Ohno, M., Ed.; Springer-Verlag: Berlin, 1990, p 467.

(38) Klymyshin, D. O.; Hromyko, O. M.; Fedorenko, V. O. Application of intergeneric conjugation of Escherichia coli-Streptomyces for transfer of recombinant DNA into S. nogalater IMET 43360 strain. *Cytology and Genetics.* **2007**, *41*, 263.

(39) Madduri, K.; Hutchinson, C. R. Functional characterization and transcriptional analysis of a gene cluster governing early and late steps in daunorubicin biosynthesis in Streptomyces peucetius. *J Bacteriol.* **1995**, *177*, 3879.

(40) Otten, S. L.; Liu, X.; Ferguson, J.; Hutchinson, C. R. Cloning and characterization of the Streptomyces peucetius dnrQS genes encoding a daunosamine biosynthesis enzyme and a glycosyl transferase involved in daunorubicin biosynthesis. *J Bacteriol.* **1995**, *177*, 6688.

(41) Otten, S. L.; Gallo, M. A.; Madduri, K.; Liu, X.; Hutchinson, C. R. Cloning and characterization of the Streptomyces peucetius dnmZUV genes encoding three enzymes required for biosynthesis of the daunorubicin precursor thymidine diphospho-L-daunosamine. *J Bacteriol.* **1997**, *179*, 4446.

(42) Madduri, K.; Kennedy, J.; Rivola, G.; Inventi-Solari, A.; Filippini, S.; Zanuso, G.; Colombo, A. L.; Gewain, K. M.; Occi, J. L.; MacNeil, D. J.; Hutchinson, C. R. Production of the antitumor drug epirubicin (4'-epidoxorubicin) and its precursor by a genetically engineered strain of Streptomyces peucetius. *Nat Biotechnol.* **1998**, *16*, 69.

(43) Siitonen, V.; Blauenburg, B.; Kallio, P.; Mantsala, P.; Metsa-Ketela, M. Discovery of a two-component monooxygenase SnoaW/SnoaL2 involved in nogalamycin biosynthesis. *Chem Biol. 19*, 638.

(44) Hutchinson, C. R. Anthracyclines. *Biotechnology.* **1995**, *28*, 331.

# References


Altenbuchner, J.; Cullum, J. DNA amplification and an unstable arginine gene in Streptomyces lividans 66. *Mol Gen Genet.* **1984**, *195*, 134.

Angel Rubio, M.; Barrado, P.; Carlos Espinosa, J.; Jimenez, A.; Fernandez Lobato, M. The pur6 gene of the puromycin biosynthetic gene cluster from Streptomyces alboniger encodes a tyrosinyl-aminonucleoside synthetase. *FEBS Lett.* **2004**, *577*, 371.

Ansari, M. Z.; Yadav, G.; Gokhale, R. S.; Mohanty, D. NRPS-PKS: a knowledge-based resource for analysis of NRPS/PKS megasynthases. *Nucleic Acids Res.* **2004**, *32*, W405.

Argoudelis, A. D.; Baczynskyj, L.; Kuo, M. T.; Laborde, A. L.; Sebek, O. K.; Truesdell, S. E.; Shilliday, F. B. Arginomycin: production, isolation, characterization and structure. *J Antibiot (Tokyo).* **1987**, *40*, 750.

Asche, C. Antitumour quinones. *Mini Rev Med Chem.* **2005**, *5*, 449.

Aziz, R. K.; Bartels, D.; Best, A. A.; DeJongh, M.; Disz, T.; Edwards, R. A.; Formsma, K.; Gerdes, S.; Glass, E. M.; Kubal, M.; Meyer, F.; Olsen, G. J.; Olson, R.; Osterman, A. L.; Overbeek, R. A.; McNeil, L. K.; Paarmann, D.; Paczian, T.; Parrello, B.; Pusch, G. D.; Reich, C.; Stevens, R.; Vassieva, O.; Vonstein, V.; Wilke, A.; Zagnitko, O. The RAST Server: rapid annotations using subsystems technology. *BMC Genomics.* **2008**, *9*, 75.

Bachmann, B. O.; Van Lanen, S. G.; Baltz, R. H. Microbial genome mining for accelerated natural products discovery: is a renaissance in the making? *J Ind Microbiol Biotechnol.* **2014**, *41*, 175.

Baltz, R. H.; Matsushima, P. Protoplast fusion in Streptomyces: conditions for efficient genetic recombination and cell regeneration. *J Gen Microbiol.* **1981**, *127*, 137.

Barrett-Lee, P. J.; Dixon, J. M.; Farrell, C.; Jones, A.; Leonard, R.; Murray, N.; Palmieri, C.; Plummer, C. J.; Stanley, A.; Verrill, M. W. Expert opinion on the use of anthracyclines in patients with advanced breast cancer at cardiac risk. *Ann Oncol.* **2009**, *20*, 816.

Bender, C. L.; Liyanage, H.; Palmer, D.; Ullrich, M.; Young, S.; Mitchell, R. Characterization of the genes controlling the biosynthesis of the polyketide phytotoxin coronatine including conjugation between coronafacic and coronamic acid. *Gene.* **1993**, *133*, 31.

Bentley, S. D.; Chater, K. F.; Cerdeno-Tarraga, A. M.; Challis, G. L.; Thomson, N. R.; James, K. D.; Harris, D. E.; Quail, M. A.; Kieser, H.; Harper, D.; Bateman, A.; Brown, S.; Chandra, G.; Chen, C. W.; Collins, M.; Cronin, A.; Fraser, A.; Goble, A.; Hidalgo, J.; Hornsby, T.; Howarth, S.; Huang, C. H.; Kieser, T.; Larke, L.;

Murphy, L.; Oliver, K.; O'Neil, S.; Rabbinowitsch, E.; Rajandream, M. A.; Rutherford, K.; Rutter, S.; Seeger, K.; Saunders, D.; Sharp, S.; Squares, R.; Squares, S.; Taylor, K.; Warren, T.; Wietzorrek, A.; Woodward, J.; Barrell, B. G.; Parkhill, J.; Hopwood, D. A. Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2). *Nature.* **2002**, *417*, 141.

Bhattacharya, K. R.; Datta, J.; Roy, D. K. A study on the Sakaguchi reaction of arginine on filter paper. *Arch Biochem Biophys.* **1958**, *77*, 297.

Bhuyan, B. K.; Blowers, C. L.; Shugars, K. D. Lethality of nogalamycin, nogalamycin analogs, and adriamycin to cells in different cell cycle phases. *Cancer Res.* **1980**, *40*, 3437.

Biermann, M.; Logan, R.; O'Brien, K.; Seno, E. T.; Nagaraja, R. R.; Schoner, B. E. Plasmid cloning vectors for the conjugal transfer of DNA from Eschericia coli to Streptomyces spp. *Gene.* **1992**, *116*, 43.

Bignell, D. R.; Seipke, R. F.; Huguet-Tapia, J. C.; Chambers, A. H.; Parry, R. J.; Loria, R. Streptomyces scabies 87-22 contains a coronafacic acid-like biosynthetic cluster that contributes to plant-microbe interactions. *Mol Plant Microbe Interact.* **2010**, *23*, 161.

Binter, A.; Oberdorfer, G.; Hofzumahaus, S.; Nerstheimer, S.; Altenbacher, G.; Gruber, K.; Macheroux, P. Characterization of the PLP-dependent aminotransferase NikK from Streptomyces tendae and its putative role in nikkomycin biosynthesis. *FEBS J.* **2011**, *278*, 4122.

Blin, K.; Medema, M. H.; Kazempour, D.; Fischbach, M. A.; Breitling, R.; Takano, E.; Weber, T. antiSMASH 2.0--a versatile platform for genome mining of secondary metabolite producers. *Nucleic Acids Res.* **2013**, *41*, W204.

Bonadonna, G.; Monfardini, S.; De Lena, M.; Fossati-Bellani, F. Clinical evaluation of adriamycin, a new antitumour antibiotic. *Br Med J.* **1969**, *3*, 503.

Borisova, S. A.; Circello, B. T.; Zhang, J. K.; van der Donk, W. A.; Metcalf, W. W. Biosynthesis of rhizocticins, antifungal phosphonate oligopeptides produced by Bacillus subtilis ATCC6633. *Chem Biol.* **2010**, *17*, 28.

Brandish, P. E.; Kimura, K. I.; Inukai, M.; Southgate, R.; Lonsdale, J. T.; Bugg, T. D. Modes of action of tunicamycin, liposidomycin B, and mureidomycin A: inhibition of phospho-N-acetylmuramyl-pentapeptide translocase from Escherichia coli. *Antimicrob Agents Chemother.* **1996**, *40*, 1640.

Bretschneider, T.; Heim, J. B.; Heine, D.; Winkler, R.; Busch, B.; Kusebauch, B.; Stehle, T.; Zocher, G.; Hertweck, C. Vinylogous chain branching catalysed by a dedicated polyketide synthase module. *Nature.* **2013**, *502*, 124.

Brown, E. G.; Konuk, M. Biosynthesis of nebularine (purine 9-B-D-ribofuranoside) involves enzymic release of hydroxylamine from adenosine. *Phytochemistry.* **1995**, *38*, 61.

Brown, T. D.; Donehower, R. C.; Grochow, L. B.; Rice, A. P.; Ettinger, D. S. A phase I study of menogaril in patients with advanced cancer. *J Clin Oncol.* **1987**, *5*, 92.

Bruntner, C.; Bormann, C. The Streptomyces tendae Tu901 L-lysine 2-aminotransferase catalyzes the initial reaction in nikkomycin D biosynthesis. *Eur J Biochem.* **1998**, *254*, 347.

Bruntner, C.; Lauer, B.; Schwarz, W.; Mohrle, V.; Bormann, C. Molecular characterization of co-transcribed genes from Streptomyces tendae Tu901 involved in the biosynthesis of the peptidyl moiety of the peptidyl nucleoside antibiotic nikkomycin. *Mol Gen Genet.* **1999**, *262*, 102.

Buja, L. M.; Ferrans, V. J.; Mayer, R. J.; Roberts, W. C.; Henderson, E. S. Cardiac ultrastructural changes induced by daunorubicin therapy. *Cancer.* **1973**, *32*, 771.

Burroughs, A. M.; Hoppe, R. W.; Goebel, N. C.; Sayyed, B. H.; Voegtline, T. J.; Schwabacher, A. W.; Zabriskie, T. M.; Silvaggi, N. R. Structural and functional characterization of MppR, an enduracididine biosynthetic enzyme from streptomyces hygroscopicus: functional diversity in the acetoacetate decarboxylase-like superfamily. *Biochemistry.* **2013**, *52*, 4492.

Cachatra, V.; Almeida, A.; Sardinha, J.; Lucas, S. D.; Gomes, A.; Vaz, P. D.; Florencio, M. H.; Nunes, R.; Vila-Vicosa, D.; Calhorda, M. J.; Rauter, A. P. Wittig Reaction: Domino Olefination and Stereoselectivity DFT Study. Synthesis of the Miharamycins' Bicyclic Sugar Moiety. *Org Lett.* **2015**, *17*, 5622.

Cai, W.; Goswami, A.; Yang, Z.; Liu, X.; Green, K. D.; Barnard-Britson, S.; Baba, S.; Funabashi, M.; Nonaka, K.; Sunkara, M.; Morris, A. J.; Spork, A. P.; Ducho, C.; Garneau-Tsodikova, S.; Thorson, J. S.; Van Lanen, S. G. The Biosynthesis of Capuramycin-type Antibiotics: IDENTIFICATION OF THE A-102395 BIOSYNTHETIC GENE CLUSTER, MECHANISM OF SELF-RESISTANCE, AND FORMATION OF URIDINE-5'-CARBOXAMIDE. *J Biol Chem.* **2015**, *290*, 13710.

Caldara, M.; Dupont, G.; Leroy, F.; Goldbeter, A.; De Vuyst, L.; Cunin, R. Arginine biosynthesis in Escherichia coli: experimental perturbation and mathematical modeling. *J Biol Chem.* **2008**, *283*, 6347.

Chan, Y. A.; Boyne, M. T., 2nd; Podevels, A. M.; Klimowicz, A. K.; Handelsman, J.; Kelleher, N. L.; Thomas, M. G. Hydroxymalonyl-acyl carrier protein (ACP) and aminomalonyl-ACP are two additional type I polyketide synthase extender units. *Proc Natl Acad Sci U S A.* **2006**, *103*, 14349.

Chan, Y. A.; Podevels, A. M.; Kevany, B. M.; Thomas, M. G. Biosynthesis of polyketide synthase extender units. *Nat Prod Rep.* **2009**, *26*, 90.

Chan, Y. A.; Thomas, M. G. Formation and characterization of acyl carrier protein-linked polyketide synthase extender units. *Methods Enzymol.* **2009**, *459*, 143.

Chan, Y. A.; Thomas, M. G. Recognition of (2S)-aminomalonyl-acyl carrier protein (ACP) and (2R)-hydroxymalonyl-ACP by acyltransferases in zwittermicin A biosynthesis. *Biochemistry.* **2010**, *49*, 3667.

Chang, C.; Zhao, L.; Yamase, H.; Liu, H. DesVI: A new member of the sugar N,N-dimethyltransferase family involved in the biosynthesis of desosamine. *Angew Chem Int Ed Engl.* **2000**, *39*, 2160.

Chatterjee, S.; Nadkarni, S. R.; Vijayakumar, E. K.; Patel, M. V.; Ganguli, B. N.; Fehlhaber, H. W.; Vertesy, L. Napsamycins, new Pseudomonas active antibiotics of the mureidomycin family from Streptomyces sp. HIL Y-82,11372. *J Antibiot (Tokyo).* **1994**, *47*, 595.

Chen, D.; Zhang, Q.; Cen, P.; Xu, Z.; Liu, W. Improvement of FK506 production in Streptomyces tsukubaensis by genetic enhancement of the supply of unusual polyketide extender units via utilization of two distinct site-specific recombination systems. *Appl Environ Microbiol.* **2012**, *78*, 5093.

Chen, H.; Du, L. Iterative polyketide biosynthesis by modular polyketide synthases in bacteria. *Appl Microbiol Biotechnol.* **2016**, *100*, 541.

Chen, H.; Guo, Z.; Liu, H.-w. Biosynthesis of yersiniose: attachment of the two-carbon branched-cain is catalyzed by a thiamine pyrophosphate-dependent flavoprotein. *J Am Chem Soc.* **1998**, *120*, 11796.

Chen, H.; Thomas, M. G.; Hubbard, B. K.; Losey, H. C.; Walsh, C. T.; Burkart, M. D. Deoxysugars in glycopeptide antibiotics: enzymatic synthesis of TDP-L-epivancosamine in chloroeremomycin biosynthesis. *Proc Natl Acad Sci U S A.* **2000**, *97*, 11942.

Chen, H.; Yamase, H.; Murakami, K.; Chang, C. W.; Zhao, L.; Zhao, Z.; Liu, H. W. Expression, purification, and characterization of two N,N-dimethyltransferases, tylM1 and desVI, involved in the biosynthesis of mycaminose and desosamine. *Biochemistry.* **2002**, *41*, 9165.

Chen, W.; Huang, T.; He, X.; Meng, Q.; You, D.; Bai, L.; Li, J.; Wu, M.; Li, R.; Xie, Z.; Zhou, H.; Zhou, X.; Tan, H.; Deng, Z. Characterization of the polyoxin biosynthetic gene cluster from Streptomyces cacaoi and engineered production of polyoxin H. *J Biol Chem.* **2009**, *284*, 10627.

Cheng, L.; Chen, W.; Zhai, L.; Xu, D.; Huang, T.; Lin, S.; Zhou, X.; Deng, Z. Identification of the gene cluster involved in muraymycin biosynthesis from Streptomyces sp. NRRL 30471. *Mol Biosyst.* **2011**, *7*, 920.

Chin, F. Y.; Leung, H. C.; Yiu, S. M. Sequence assembly using next generation sequencing data--challenges and solutions. *Sci China Life Sci.* **2014**, *57*, 1140.

Cock, P. J. A.; Fields, C. J.; Goto, N.; Heuer, M. L.; Rice, P. M. The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants. *Nucleic Acids Res.* **2010**, *38*, 1767.

Comin, M.; Schimd, M. Assembly-free genome comparison based on next-generation sequencing reads and variable length patterns. *BMC Bioinformatics.* **2014**, *15 Suppl 9*, S1.

Compeau, P. E.; Pevzner, P. A.; Tesler, G. How to apply de Bruijn graphs to genome assembly. *Nat Biotechnol.* **2011**, *29*, 987.

Cone, M. C.; Petrich, A. K.; Gould, S. J.; Zabriskie, T. M. Cloning and heterologous expression of blasticidin S biosynthetic genes from Streptomyces griseochromogenes. *J Antibiot (Tokyo).* **1998**, *51*, 570.

Cone, M. C.; Yin, X.; Grochowski, L. L.; Parker, M. R.; Zabriskie, T. M. The blasticidin S biosynthesis gene cluster from Streptomyces griseochromogenes: sequence analysis, organization, and initial characterization. *Chembiochem.* **2003**, *4*, 821.

Cooper, R.; Horan, A. C.; Gunnarsson, I.; Patel, M.; Truumees, I. Nebularine from a novel Microbispora sp. *J Ind Microbiol* **1986**, *1*, 275.

Corman, J.; Tsuchiya, H. M.; Koepsell, H. J.; Benedict, R. G.; Kelley, S. E.; Feger, V. H.; Dworschack, R. G.; Jackson, R. W. Oxygen absorption rates in laboratory and pilot plant equipment. *Appl Microbiol.* **1957**, *5*, 313.

Coutinho, P. M.; Deleury, E.; Davies, G. J.; Henrissat, B. An evolving hierarchical family classification for glycosyltransferases. *J Mol Biol.* **2003**, *328*, 307.

Cragg, G. M.; Newman, D. J. Natural products: a continuing source of novel drug leads. *Biochim Biophys Acta.* **2013**, *1830*, 3670.

Darling, A. C.; Mau, B.; Blattner, F. R.; Perna, N. T. Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res.* **2004**, *14*, 1394.

Darling, A. E.; Treangen, T. J.; Messeguer, X.; Perna, N. T. Analyzing patterns of microbial evolution using the mauve genome alignment system. *Methods Mol Biol.* **2007**, *396*, 135.

Darling, A. E.; Tritt, A.; Eisen, J. A.; Facciotti, M. T. Mauve assembly metrics. *Bioinformatics.* **2011**, *27*, 2756.

DaVanzo, J. P.; Matthews, R. J.; Stafford, J. E. Studies on the mechanism of action of aminooxyacetic acid. I. Reversal of aminooxyacetic acid-induced convulsions by various agents. *Toxicology and Applied Pharmacology.* **1964**, *6*, 388.

Davies, K. J.; Doroshow, J. H. Redox cycling of anthracyclines by cardiac mitochondria. I. Anthracycline radical formation by NADH dehydrogenase. *J Biol Chem.* **1986**, *261*, 3060.

Dawlaty, J.; Zhang, X.; Fischbach, M. A.; Clardy, J. Dapdiamides, tripeptide antibiotics formed by unconventional amide ligases. *J Nat Prod.* **2010**, *73*, 441.

Dekleva, M. L.; Titus, J. A.; Strohl, W. R. Nutrient effects on anthracycline production by Streptomyces peucetius in a defined medium. *Can J Microbiol.* **1985**, *31*, 287.

Di Marco, A.; Casazza, A. M.; Gambetta, R.; Supino, R.; Zunino, F. Relationship between activity and amino sugar stereochemistry of daunorubicin and adriamycin derivatives. *Cancer Res.* **1976**, *36*, 1962.

Dias, D. A.; Urban, S.; Roessner, U. A historical overview of natural products in drug discovery. *Metabolites.* **2012**, *2*, 303.

Dimarco, A.; Gaetani, M.; Dorigotti, L.; Soldati, M.; Bellini, O. Daunomycin: A New Antibiotic with Antitumor Activity. *Cancer Chemother Rep.* **1964**, *38*, 31.

Dimarco, A.; Gaetani, M.; Orezzi, P.; Scarpinato, B. M.; Silvestrini, R.; Soldati, M.; Dasdia, T.; Valentini, L. 'Daunomycin', a New Antibiotic of the Rhodomycin Group. *Nature.* **1964**, *201*, 706.

Dimarco, A.; Soldati, M.; Fioretti, A.; Dasdia, T. Activity of Daunomycin, a New Antitumor Antibiotic, on Normal and Neoplastic Cells Grown in Vitro. *Cancer Chemother Rep.* **1964**, *38*, 39.

Dobashi, K.; Matsuda, N.; Hamada, M.; Naganawa, H.; Takita, T.; Takeuchi, T. Novel antifungal antibiotics octacosamicins A and B. I. Taxonomy, fermentation and isolation, physico-chemical properties and biological activities. *J Antibiot (Tokyo).* **1988**, *41*, 1525.

Dobashi, K.; Naganawa, H.; Takahashi, Y.; Takita, T.; Takeuchi, T. Novel antifungal antibiotics octacosamicins A and B. II. The structure elucidation using various NMR spectroscopic methods. *J Antibiot (Tokyo).* **1988**, *41*, 1533.

Dong, C.; Major, L. L.; Srikannathasan, V.; Errey, J. C.; Giraud, M. F.; Lam, J. S.; Graninger, M.; Messner, P.; McNeil, M. R.; Field, R. A.; Whitfield, C.; Naismith, J. H. RmlC, a C3' and C5' carbohydrate epimerase, appears to operate via an intermediate with an unusual twist boat conformation. *J Mol Biol.* **2007**, *365*, 146.

Dong, C.; Major, L. L.; Allen, A.; Blankenfeldt, W.; Maskell, D.; Naismith, J. H. High-resolution structures of RmlC from Streptococcus suis in complex with substrate analogs locate the active site of this class of enzyme. *Structure.* **2003**, *11*, 715.

Doroshow, J. H.; Davies, K. J. Redox cycling of anthracyclines by cardiac mitochondria. II. Formation of superoxide anion, hydrogen peroxide, and hydroxyl radical. *J Biol Chem.* **1986**, *261*, 3068.

Du, L.; Lou, L. PKS and NRPS release mechanisms. *Nat Prod Rep.* **2010**, *27*, 255.

Du, W.; Huang, D.; Xia, M.; Wen, J.; Huang, M. Improved FK506 production by the precursors and product-tolerant mutant of Streptomyces tsukubaensis based on genome shuffling and dynamic fed-batch strategies. *J Ind Microbiol Biotechnol.* **2014**, *41*, 1131.

Elander, R. P. Industrial production of beta-lactam antibiotics. *Appl Microbiol Biotechnol.* **2003**, *61*, 385.

Fawaz, M. V.; Topper, M. E.; Firestine, S. M. The ATP-grasp enzymes. *Bioorg Chem.* **2011**, *39*, 185.

Feng, J.; Wu, J.; Dai, N.; Lin, S.; Xu, H. H.; Deng, Z.; He, X. Discovery and characterization of BlsE, a radical S-adenosyl-L-methionine decarboxylase involved in the blasticidin S biosynthetic pathway. *PLoS One.* **2013**, *8*, e68545.

Feng, J.; Wu, J.; Gao, J.; Xia, Z.; Deng, Z.; He, X. Biosynthesis of the beta-methylarginine residue of peptidyl nucleoside arginomycin in Streptomyces arginensis NRRL 15941. *Appl Environ Microbiol.* **2014**, *80*, 5021.

Fernandez-Moreno, M. A.; Vallin, C.; Malpartida, F. Streptothricin biosynthesis is catalyzed by enzymes related to nonribosomal peptide bond formation. *J Bacteriol.* **1997**, *179*, 6929.

Fischer, B.; Keller-Schierlein, W.; Kneifel, H.; Konig, W. A.; Loeffler, W.; Muller, A.; Muntwyler, R.; Zahner, H. Metabolic products of microorganisms. 118. Delta-N-hydroxy-L-arginine, an amino acid antagonist from Nannizzia gypsea. *Archiv fur Mikrobiologie.* **1973**, *93*, 203.

Fischer, K.; Llamas, A.; Tejada-Jimenez, M.; Schrader, N.; Kuper, J.; Ataya, F. S.; Galvan, A.; Mendel, R. R.; Fernandez, E.; Schwarz, G. Function and structure of the molybdenum cofactor carrier protein from Chlamydomonas reinhardtii. *J Biol Chem.* **2006**, *281*, 30186.

Frederick, C. A.; Williams, L. D.; Ughetto, G.; van der Marel, G. A.; van Boom, J. H.; Rich, A.; Wang, A. H. Structural comparison of anticancer drug-DNA complexes: adriamycin and daunomycin. *Biochemistry.* **1990**, *29*, 2538.

Funabashi, M.; Baba, S.; Nonaka, K.; Hosobuchi, M.; Fujita, Y.; Shibata, T.; Van Lanen, S. G. The biosynthesis of liposidomycin-like A-90289 antibiotics featuring a new type of sulfotransferase. *Chembiochem.* **2010**, *11*, 184.

Funabashi, M.; Nonaka, K.; Yada, C.; Hosobuchi, M.; Masuda, N.; Shibata, T.; Van Lanen, S. G. Identification of the biosynthetic gene cluster of A-500359s in Streptomyces griseus SANK60196. *J Antibiot (Tokyo).* **2009**, *62*, 325.

Funabashi, M.; Yang, Z.; Nonaka, K.; Hosobuchi, M.; Fujita, Y.; Shibata, T.; Chi, X.; Van Lanen, S. G. An ATP-independent strategy for amide bond formation in antibiotic biosynthesis. *Nat Chem Biol.* **2010**, *6*, 581.

Galperin, M. Y.; Koonin, E. V. From complete genome sequence to 'complete' understanding? *Trends Biotechnol.* **2010**, *28*, 398.

Gelfand, M. S.; Cleveland, K. O. Ceftolozane/Tazobactam Therapy of Respiratory Infections due to Multidrug-Resistant Pseudomonas aeruginosa. *Clin Infect Dis.* **2015**, *61*, 853.

Genilloud, O. The re-emerging role of microbial natural products in antibiotic discovery. *Antonie Van Leeuwenhoek.* **2014**, *106*, 173.

Genilloud, O.; Gonzalez, I.; Salazar, O.; Martin, J.; Tormo, J. R.; Vicente, F. Current approaches to exploit actinomycetes as a source of novel natural products. *J Ind Microbiol Biotechnol.* **2011**, *38*, 375.

Gewirtz, D. A. A critical evaluation of the mechanisms of action proposed for the antitumor effects of the anthracycline antibiotics adriamycin and daunorubicin. *Biochem Pharmacol.* **1999**, *57*, 727.

Goodman, M. F.; Hopkins, R. L.; Lasken, R.; Mhaskar, D. N. The biochemical basis of 5-bromouracil- and 2-aminopurine-induced mutagenesis. *Basic Life Sci.* **1985**, *31*, 409.

Gordon, A. FASTX-Toolkit. Computer program distributed by the author. **2011**.

Goswami, A.; Van Lanen, S. G. Enzymatic strategies and biocatalysts for amide bond formation: tricks of the trade outside of the ribosome. *Mol Biosyst.* **2015**, *11*, 338.

Goto, T.; Toya, Y.; Kondo, T. Structure of amipurimycin, a new nucleoside antibiotic produced by Streptomyces novoguineensis. *Nucleic Acids Symp Ser.* **1980**, s73.

Gould, S. J.; Guo, J. Cytosylglucuronic acid synthase (cytosine: UDP-glucuronosyltransferase) from Streptomyces griseochromogenes, the first prokaryotic UDP-glucuronosyltransferase. *J Bacteriol.* **1994**, *176*, 1282.

Gould, S. J.; Guo, J.; Geitmann, A.; DeJesus, K. Nucleoside intermediates in blasticidin S biosynthesis identified by the in vivo use of enzyme inhibitors. *Canadian Journal of Chemistry.* **1994**, *72*, 6.

Graninger, M.; Nidetzky, B.; Heinrichs, D. E.; Whitfield, C.; Messner, P. Characterization of dTDP-4-dehydrorhamnose 3,5-epimerase and dTDP-4-dehydrorhamnose reductase, required for dTDP-L-rhamnose biosynthesis in Salmonella enterica serovar Typhimurium LT2. *J Biol Chem.* **1999**, *274*, 25069.

Griffith, O. W.; Stuehr, D. J. Nitric oxide synthases: properties and catalytic mechanism. *Annu Rev Physiol.* **1995**, *57*, 707.

Gunasekera, A.; Alvarez, F. J.; Douglas, L. M.; Wang, H. X.; Rosebrock, A. P.; Konopka, J. B. Identification of GIG1, a GlcNAc-induced gene in Candida albicans needed for normal sensitivity to the chitin synthase inhibitor nikkomycin Z. *Eukaryot Cell.* **2010**, *9*, 1476.

Guo, Z.; Li, J.; Qin, H.; Wang, M.; Lv, X.; Li, X.; Chen, Y. Biosynthesis of the carbamoylated D-gulosamine moiety of streptothricins: involvement of a guanidino-N-glycosyltransferase and an N-acetyl-D-gulosamine deacetylase. *Angew Chem Int Ed Engl.* **2015**, *54*, 5175.

Han, A. R.; Park, J. W.; Lee, M. K.; Ban, Y. H.; Yoo, Y. J.; Kim, E. J.; Kim, E.; Kim, B. G.; Sohng, J. K.; Yoon, Y. J. Development of a Streptomyces venezuelae-based combinatorial biosynthetic system for the production of glycosylated derivatives of doxorubicin and its biosynthetic intermediates. *Appl Environ Microbiol.* **2011**, *77*, 4912.

Han, L.; Schwabacher, A. W.; Moran, G. R.; Silvaggi, N. R. Streptomyces wadayamensis MppP Is a Pyridoxal 5'-Phosphate-Dependent l-Arginine alpha-Deaminase, gamma-Hydroxylase in the Enduracididine Biosynthetic Pathway. *Biochemistry.* **2015**, *54*, 7029.

Hanessian, S.; Huang, G.; Chenel, C.; Machaalani, R.; Loiseleur, O. Total synthesis of N-malayamycin A and related bicyclic purine and pyrimidine nucleosides. *J Org Chem.* **2005**, *70*, 6721.

Hanessian, S.; Marcotte, S.; Machaalani, R.; Huang, G. Total synthesis and structural confirmation of malayamycin A: a novel bicyclic C-nucleoside from Streptomyces malaysiensis. *Org Lett.* **2003**, *5*, 4277.

Harada, S.; Kishi, T. Isolation and characterization of a new nucleoside antibiotic, amipurimycin. *J Antibiot (Tokyo).* **1977**, *30*, 11.

Haupt, I.; Hubener, R.; Thrum, H. Streptothricin F, an inhibitor of protein synthesis with miscoding activity. *J Antibiot (Tokyo).* **1978**, *31*, 1137.

Hauser, F. M.; Ganguly, D. Synthetic studies on nogarol anthracyclines. Enantioselective total synthesis of an aminohydroxy epoxybenzoxocin. *J Org Chem.* **2000**, *65*, 1842.

Heine, D.; Bretschneider, T.; Sundaram, S.; Hertweck, C. Enzymatic polyketide chain branching to give substituted lactone, lactam, and glutarimide heterocycles. *Angew Chem Int Ed Engl.* **2014**, *53*, 11645.

Heine, D.; Sundaram, S.; Bretschneider, T.; Hertweck, C. Twofold polyketide branching by a stereoselective enzymatic Michael addition. *Chem Commun (Camb).* **2015**, *51*, 9872.

Henrich, C. J.; Beutler, J. A. Matching the power of high throughput screening to the chemical diversity of natural products. *Nat Prod Rep.* **2013**, *30*, 1284.

Henry, D. W. Structure-activity relationships among daunorubicin and adriamycin analogs. *Cancer Treat Rep.* **1979**, *63*, 845.

Hertweck, C. The biosynthetic logic of polyketide diversity. *Angew Chem Int Ed Engl.* **2009**, *48*, 4688.

Hertweck, C. Decoding and reprogramming complex polyketide assembly lines: prospects for synthetic biology. *Trends Biochem Sci.* **2015**, *40*, 189.

Hofmann, C.; Boll, R.; Heitmann, B.; Hauser, G.; Durr, C.; Frerich, A.; Weitnauer, G.; Glaser, S. J.; Bechthold, A. Genes encoding enzymes responsible for biosynthesis of L-lyxose and attachment of eurekanate during avilamycin biosynthesis. *Chem Biol.* **2005**, *12*, 1137.

Hong, L.; Zhao, Z.; Melancon, C. E., 3rd; Zhang, H.; Liu, H. W. In vitro characterization of the enzymes involved in TDP-D-forosamine biosynthesis in the spinosyn pathway of Saccharopolyspora spinosa. *J Am Chem Soc.* **2008**, *130*, 4954.

Hopkins, R. L.; Goodman, M. F. Ribonucleoside and deoxyribonucleoside triphosphate pools during 2-aminopurine mutagenesis in T4 mutator-, wild type-, and antimutator-infected Escherichia coli. *J Biol Chem.* **1985**, *260*, 6618.

Houge-Frydrych, C. S.; Readshaw, S. A.; Bell, D. J. SB-219383, a novel tyrosyl tRNA synthetase inhibitor from a Micromonospora sp. II. Structure determination. *J Antibiot (Tokyo).* **2000**, *53*, 351.

Hutchinson, C. R. Anthracyclines. *Biotechnology.* **1995**, *28*, 331.

Illumina: 2012.

Isono, F.; Inukai, M. Mureidomycin A, a new inhibitor of bacterial peptidoglycan synthesis. *Antimicrob Agents Chemother.* **1991**, *35*, 234.

Isono, F.; Sakaida, Y.; Takahashi, S.; Kinoshita, T.; Nakamura, T.; Inukai, M. Mureidomycins E and F, minor components of mureidomycins. *J Antibiot (Tokyo).* **1993**, *46*, 1203.

Isono, K. Nucleoside antibiotics: structure, biological activity, and biosynthesis. *J Antibiot (Tokyo).* **1988**, *41*, 1711.

Isono, K. Current progress on nucleoside antibiotics. *Pharmacol Ther.* **1991**, *52*, 269.

Iwamoto, T.; Tsujii, E.; Ezaki, M.; Fujie, A.; Hashimoto, S.; Okuhara, M.; Kohsaka, M.; Imanaka, H.; Kawabata, K.; Inamoto, Y.; et al. FR109615, a new antifungal antibiotic from Streptomyces setonii. Taxonomy, fermentation, isolation, physico-chemical properties and biological activity. *J Antibiot (Tokyo).* **1990**, *43*, 1.

Iwasa, T.; Kishi, T.; Matsuura, K.; Wakae, O. Streptomyces novoguineensis sp. Nov., an amipurimycin producer, and antimicrobial activity of amipurimycin. *J Antibiot (Tokyo).* **1977**, *30*, 1.

Iyer, L. M.; Abhiman, S.; Maxwell Burroughs, A.; Aravind, L. Amidoligases with ATP-grasp, glutamine synthetase-like and acetyltransferase-like domains: synthesis of novel metabolites and peptide modifications of proteins. *Mol Biosyst.* **2009**, *5*, 1636.

Jain, K. K.; Casper, E. S.; Geller, N. L.; Hakes, T. B.; Kaufman, R. J.; Currie, V.; Schwartz, W.; Cassidy, C.; Petroni, G. R.; Young, C. W.; et al. A prospective randomized comparison of epirubicin and doxorubicin in patients with advanced breast cancer. *J Clin Oncol.* **1985**, *3*, 818.

Jannazzo, A.; Hoffman, J.; Lutz, M. Monitoring of anthracycline-induced cardiotoxicity. *Ann Pharmacother.* **2008**, *42*, 99.

Jiang, L.; Wei, J.; Li, L.; Niu, G.; Tan, H. Combined gene cluster engineering and precursor feeding to improve gougerotin production in Streptomyces graminearus. *Appl Microbiol Biotechnol.* **2013**, *97*, 10469.

Jorpes, E.; Thoren, S. The use of the Sakaguchi reaction for the quantitative determination of arginine. *Biochem J.* **1932**, *26*, 1504.

Ju, K. S.; Gao, J.; Doroghazi, J. R.; Wang, K. K.; Thibodeaux, C. J.; Li, S.; Metzger, E.; Fudala, J.; Su, J.; Zhang, J. K.; Lee, J.; Cioni, J. P.; Evans, B. S.; Hirota, R.; Labeda, D. P.; van der Donk, W. A.; Metcalf, W. W. Discovery of phosphonic acid natural products by mining the genomes of 10,000 actinomycetes. *Proc Natl Acad Sci U S A.* **2015**, *112*, 12175.

Katane, M.; Kawata, T.; Nakayama, K.; Saitoh, Y.; Kaneko, Y.; Matsuda, S.; Miyamoto, T.; Sekine, M.; Homma, H. Characterization of the enzymatic and structural properties of human D-aspartate oxidase and comparison with those of the rat and mouse enzymes. *Biol Pharm Bull.* **2015**, *38*, 298.

Katz, L.; Baltz, R. H. Natural product discovery: past, present, and future. *J Ind Microbiol Biotechnol.* **2016**.

Kaye, S.; Merry, S. Tumour cell resistance to anthracyclines--a review. *Cancer Chemother Pharmacol.* **1985**, *14*, 96.

Kaysser, L.; Siebenberg, S.; Kammerer, B.; Gust, B. Analysis of the liposidomycin gene cluster leads to the identification of new caprazamycin derivatives. *Chembiochem.* **2010**, *11*, 191.

Kaysser, L.; Tang, X.; Wemakor, E.; Sedding, K.; Hennig, S.; Siebenberg, S.; Gust, B. Identification of a napsamycin biosynthesis gene cluster by genome mining. *Chembiochem.* **2011**, *12*, 477.

Kevany, B. M.; Rasko, D. A.; Thomas, M. G. Characterization of the complete zwittermicin A biosynthesis gene cluster from Bacillus cereus. *Appl Environ Microbiol.* **2009**, *75*, 1144.

Khosla, C. Quo vadis, enzymology? *Nat Chem Biol.* **2015**, *11*, 438.

Kieser, T.; Bibb, M. J.; Buttner, M. J.; Chater, K. F.; Hopwood, D. A. *Practical streptomyces genetics*; The John Innes Foundation: Norwich, U.K., 2000.

Kim, E.; Moore, B. S.; Yoon, Y. J. Reinvigorating natural product combinatorial biosynthesis with synthetic biology. *Nat Chem Biol.* **2015**, *11*, 649.

Kino, K.; Kotanaka, Y.; Arai, T.; Yagasaki, M. A novel L-amino acid ligase from Bacillus subtilis NBRC3134, a microorganism producing peptide-antibiotic rhizocticin. *Biosci Biotechnol Biochem.* **2009**, *73*, 901.

Kino, K.; Noguchi, A.; Nakazawa, Y.; Yagasaki, M. A novel l-amino acid ligase from bacillus licheniformis. *J Biosci Bioeng.* **2008**, *106*, 313.

Kisner, D. L.; Catane, R.; Muggia, F. M. In *Cancer Chemo- and Immunopharmacology. 1. Chemopharmacology*; Mathe, G., Muggia, F. M., Eds.; Springer Berlin Heidelberg: Berlin, 1980; Vol. 74, p 258.

Kleczkowski, L. A. Inhibitors of photosynthetic enzymes/ carriers and metabolism. *Annual Review of Plant Physiology and Plant Molecular Biology.* **1994**, *45*, 339.

Kleczkowski, L. A.; Randall, D. D.; Blevins, D. G. Inhibition of Spinach Leaf NADPH(NADH)-Glyoxylate Reductase by Acetohydroxamate, Aminooxyacetate, and Glycidate. *Plant Physiol.* **1987**, *84*, 619.

Kleiger, G.; Eisenberg, D. GXXXG and GXXXA motifs stabilize FAD and NAD(P)-binding Rossmann folds through C(alpha)-H... O hydrogen bonds and van der waals interactions. *J Mol Biol.* **2002**, *323*, 69.

Klementz, D.; Doring, K.; Lucas, X.; Telukunta, K. K.; Erxleben, A.; Deubel, D.; Erber, A.; Santillana, I.; Thomas, O. S.; Bechthold, A.; Gunther, S. StreptomeDB 2.0-an extended resource of natural products produced by streptomycetes. *Nucleic Acids Res.* **2015**.

Klymyshin, D. O.; Hromyko, O. M.; Fedorenko, V. O. Application of intergeneric conjugation of Escherichia coli-Streptomyces for transfer of recombinant DNA into S. nogalater IMET 43360 strain. *Cytology and Genetics.* **2007**, *41*, 263.

Knapp, S.; Morriello, G. J.; Nandan, S. R.; Emge, T. J.; Doss, G. A.; Mosley, R. T.; Chen, L. Assignment of the liposidomycin diazepanone stereochemistry. *J Org Chem.* **2001**, *66*, 5822.

Konishi, M.; Nishio, M.; Saitoh, K.; Miyaki, T.; Oki, T.; Kawaguchi, H. Cispentacin, a new antifungal antibiotic. I. Production, isolation, physico-chemical properties and structure. *J Antibiot (Tokyo).* **1989**, *42*, 1749.

Kothe, M.; Powers-Lee, S. G. Nucleotide recognition in the ATP-grasp protein carbamoyl phosphate synthetase. *Protein Sci.* **2004**, *13*, 466.

Kuroha, T.; Tokunaga, H.; Kojima, M.; Ueda, N.; Ishida, T.; Nagawa, S.; Fukuda, H.; Sugimoto, K.; Sakakibara, H. Functional analyses of LONELY GUY cytokinin-activating enzymes reveal the importance of the direct activation pathway in Arabidopsis. *Plant Cell.* **2009**, *21*, 3152.

Kusebauch, B.; Busch, B.; Scherlach, K.; Roth, M.; Hertweck, C. Polyketide-chain branching by an enzymatic Michael addition. *Angew Chem Int Ed Engl.* **2009**, *48*, 5001.

Laatsch, H.; Fotso, S. Naturally Occurring Anthracyclines. *Topics in Current Chemistry.* **2008**, *282*, 3.

Lacalle, R. A.; Tercero, J. A.; Jimenez, A. Cloning of the complete biosynthetic gene cluster for an aminonucleoside antibiotic, puromycin, and its regulated expression in heterologous hosts. *EMBO J.* **1992**, *11*, 785.

Lairson, L. L.; Henrissat, B.; Davies, G. J.; Withers, S. G. Glycosyltransferases: structures, functions, and mechanisms. *Annu Rev Biochem.* **2008**, *77*, 521.

Lauer, B.; Russwurm, R.; Bormann, C. Molecular characterization of two genes from Streptomyces tendae Tu901 required for the formation of the 4-formyl-4-imidazolin-2-one-containing nucleoside moiety of the peptidyl nucleoside antibiotic nikkomycin. *Eur J Biochem.* **2000**, *267*, 1698.

Lauer, B.; Russwurm, R.; Schwarz, W.; Kalmanczhelyi, A.; Bruntner, C.; Rosemeier, A.; Bormann, C. Molecular characterization of co-transcribed genes from Streptomyces tendae Tu901 involved in the biosynthesis of the peptidyl moiety and assembly of the peptidyl nucleoside antibiotic nikkomycin. *Mol Gen Genet.* **2001**, *264*, 662.

Lawrence, J. Selfish operons: the evolutionary impact of gene clustering in prokaryotes and eukaryotes. *Curr Opin Genet Dev.* **1999**, *9*, 642.

Lawrence, J. G. Gene organization: selection, selfishness, and serendipity. *Annu Rev Microbiol.* **2003**, *57*, 419.

Layer, G.; Gaddam, S. A.; Ayala-Castro, C. N.; Ollagnier-de Choudens, S.; Lascoux, D.; Fontecave, M.; Outten, F. W. SufE transfers sulfur from SufS to SufB for iron-sulfur cluster assembly. *J Biol Chem.* **2007**, *282*, 13342.

Lee, H.; Tang, H. Next-generation sequencing technologies and fragment assembly algorithms. *Methods Mol Biol.* **2012**, *855*, 155.

Lee, T. V.; Johnson, R. D.; Arcus, V. L.; Lott, J. S. Prediction of the substrate for nonribosomal peptide synthetase (NRPS) adenylation domains by virtual screening. *Proteins.* **2015**, *83*, 2052.

Li, L.; Xu, Z.; Xu, X.; Wu, J.; Zhang, Y.; He, X.; Zabriskie, T. M.; Deng, Z. The mildiomycin biosynthesis: initial steps for sequential generation of 5-hydroxymethylcytidine 5'-monophosphate and 5-hydroxymethylcytosine in Streptoverticillium rimofaciens ZJU5119. *Chembiochem.* **2008**, *9*, 1286.

Li, W.; Csukai, M.; Corran, A.; Crowley, P.; Solomon, P. S.; Oliver, R. P. Malayamycin, a new streptomycete antifungal compound, specifically inhibits sporulation of

Stagonospora nodorum (Berk) Castell and Germano, the cause of wheat glume blotch disease. *Pest Manag Sci.* **2008**, *64*, 1294.

Li, Y.; Zeng, H.; Tan, H. Cloning, function, and expression of sanS: a gene essential for nikkomycin biosynthesis of Streptomyces ansochromogenes. *Curr Microbiol.* **2004**, *49*, 128.

Liao, G.; Li, J.; Li, L.; Yang, H.; Tian, Y.; Tan, H. Cloning, reassembling and integration of the entire nikkomycin biosynthetic gene cluster into Streptomyces ansochromogenes lead to an improved nikkomycin production. *Microb Cell Fact.* **2010**, *9*, 6.

Liaw, Y. C.; Gao, Y. G.; Robinson, H.; van der Marel, G. A.; van Boom, J. H.; Wang, A. H. Antitumor drug nogalamycin binds DNA in both grooves simultaneously: molecular structure of nogalamycin-DNA complex. *Biochemistry.* **1989**, *28*, 9913.

Lin, C. I.; McCarty, R. M.; Liu, H. W. The biosynthesis of nitrogen-, sulfur-, and high-carbon chain-containing sugars. *Chem Soc Rev.* **2013**, *42*, 4377.

Liou, G. F.; Khosla, C. Building-block selectivity of polyketide synthases. *Curr Opin Chem Biol.* **2003**, *7*, 279.

Liu, W.; Peterson, P. E.; Carter, R. J.; Zhou, X.; Langston, J. A.; Fisher, A. J.; Toney, M. D. Crystal structures of unbound and aminooxyacetate-bound Escherichia coli gamma-aminobutyrate aminotransferase. *Biochemistry.* **2004**, *43*, 10896.

Loscher, W.; Honack, D.; Gramer, M. Use of inhibitors of gamma-aminobutyric acid (GABA) transaminase for the estimation of GABA turnover in various brain regions of rats: a reevaluation of aminooxyacetic acid. *J Neurochem.* **1989**, *53*, 1737.

Lown, J. W. The mechanism of action of quinone antibiotics. *Mol Cell Biochem.* **1983**, *55*, 17.

Lucas, X.; Senger, C.; Erxleben, A.; Gruning, B. A.; Doring, K.; Mosch, J.; Flemming, S.; Gunther, S. StreptomeDB: a resource for natural compounds isolated from Streptomyces species. *Nucleic Acids Res.* **2013**, *41*, D1130.

Luo, Y.; Ruan, L. F.; Zhao, C. M.; Wang, C. X.; Peng, D. H.; Sun, M. Validation of the intact zwittermicin A biosynthetic gene cluster and discovery of a complementary resistance mechanism in Bacillus thuringiensis. *Antimicrob Agents Chemother.* **2011**, *55*, 4161.

MacNeil, D. J.; Gewain, K. M.; Ruby, C. L.; Dezeny, G.; Gibbons, P. H.; MacNeil, T. Analysis of Streptomyces avermitilis genes required for avermectin biosynthesis utilizing a novel integration vector. *Gene.* **1992**, *111*, 61.

Madduri, K.; Hutchinson, C. R. Functional characterization and transcriptional analysis of a gene cluster governing early and late steps in daunorubicin biosynthesis in Streptomyces peucetius. *J Bacteriol.* **1995**, *177*, 3879.

Madduri, K.; Kennedy, J.; Rivola, G.; Inventi-Solari, A.; Filippini, S.; Zanuso, G.; Colombo, A. L.; Gewain, K. M.; Occi, J. L.; MacNeil, D. J.; Hutchinson, C. R. Production of the antitumor drug epirubicin (4'-epidoxorubicin) and its precursor by a genetically engineered strain of Streptomyces peucetius. *Nat Biotechnol.* **1998**, *16*, 69.

Madduri, K.; Waldron, C.; Merlo, D. J. Rhamnose biosynthesis pathway supplies precursors for primary and secondary metabolism in Saccharopolyspora spinosa. *J Bacteriol.* **2001**, *183*, 5632.

Maehr, H.; Blount, J. F.; Pruess, D. L.; Yarmchuk, L.; Kellett, M. Antimetabolite produced by microorganisms. VIII. N5-hydroxy-L-arginine, a new naturally occurring amino acid. *J. Antibiot.* **1973**, *26*, 284.

Makarieva, T. N.; Ogurtsova, E. K.; Denisenko, V. A.; Dmitrenok, P. S.; Tabakmakher, K. M.; Guzii, A. G.; Pislyagin, E. A.; Es'kov, A. A.; Kozhemyako, V. B.; Aminin, D. L.; Wang, Y. M.; Stonik, V. A. Urupocidin A: a new, inducing iNOS expression bicyclic guanidine alkaloid from the marine sponge Monanchora pulchra. *Org Lett.* **2014**, *16*, 4292.

Makarova, K. S.; Mironov, A. A.; Gelfand, M. S. Conservation of the binding site for the arginine repressor in all bacterial lineages. *Genome Biol.* **2001**, *2*, RESEARCH0013.

Mansoorabadi, S. O.; Thibodeaux, C. J.; Liu, H. W. The diverse roles of flavin coenzymes--nature's most versatile thespians. *J Org Chem.* **2007**, *72*, 6329.

Marahiel, M. A. A structural model for multimodular NRPS assembly lines. *Nat Prod Rep.* **2015**.

Marcelo, F.; Jimenez-Barbero, J.; Marrot, J.; Rauter, A. P.; Sinay, P.; Bleriot, Y. Stereochemical assignment and first synthesis of the core of miharamycin antibiotics. *Chemistry.* **2008**, *14*, 10066.

Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal.* **2011**, *17*, 10.

Maruyama, C.; Toyoda, J.; Kato, Y.; Izumikawa, M.; Takagi, M.; Shin-ya, K.; Katano, H.; Utagawa, T.; Hamano, Y. A stand-alone adenylation domain forms amide bonds in streptothricin biosynthesis. *Nat Chem Biol.* **2012**, *8*, 791.

Mazodier, P.; Petter, R.; Thompson, C. Intergeneric conjugation between Escherichia coli and Streptomyces species. *J Bacteriol.* **1989**, *171*, 3583.

McCarthy, K. Pseudomonas aeruginosa: evolution of antimicrobial resistance and implications for therapy. *Semin Respir Crit Care Med.* **2015**, *36*, 44.

McDaniel, L. E.; Bailey, E. G. Effect of shaking speed and type of closure on shake flask cultures. *Appl Microbiol.* **1969**, *17*, 286.

McDonald, L. A.; Barbieri, L. R.; Carter, G. T.; Lenoy, E.; Lotvin, J.; Petersen, P. J.; Siegel, M. M.; Singh, G.; Williamson, R. T. Structures of the muraymycins, novel peptidoglycan biosynthesis inhibitors. *J Am Chem Soc.* **2002**, *124*, 10260.

Metsa-Ketela, M.; Niemi, J.; Mantsala, P.; Schneider, G. Anthracycline Biosynthesis: Genes, Enzymes and Mechanisms. *Topics in Current Chemistry.* **2008**, *282*, 101.

Misra, A. The use of antibiotics for the control of plant virus diseases. *Journal of Plant Diseases and Protection.* **1977**, *84*, 244.

Miyashiro, S.; Ando, T.; Hirayama, K.; Kida, T.; Shibai, H.; Murai, A.; Shiio, T.; Udaka, S. New streptothricin-group antibiotics, AN-201 I and II. Screening, fermentation, isolation, structure and biological activity. *J Antibiot (Tokyo).* **1983**, *36*, 1638.

Nemoto, A.; Hoshino, Y.; Yazawa, K.; Ando, A.; Mikami, Y.; Komaki, H.; Tanaka, Y.; Grafe, U. Asterobactin, a new siderophore group antibiotic from Nocardia asteroides. *J. Antibiot.* **2002**, *55*, 593.

Newman, D. J.; Cragg, G. M. Natural products as sources of new drugs over the 30 years from 1981 to 2010. *J Nat Prod.* **2012**, *75*, 311.

Niemi, J.; Metsa-Ketela, M.; Schneider, G.; Mantsala, P. Biosynthetic Anthracycline Variants. *Topics in Current Chemistry.* **2008**, *282*, 75.

Niida, T.; Yumoto, H.; Tsuruoka, T.; Hamamoto, K.; Shomura, J.; Ohashi, T.; Antibiotics Obtained from Streptomyces Miharaensis; U.S. Patent 3,678,159, July 18, 1972.

Niu, G.; Li, L.; Wei, J.; Tan, H. Cloning, heterologous expression, and characterization of the gene cluster required for gougerotin biosynthesis. *Chem Biol.* **2013**, *20*, 34.

Niu, G.; Tan, H. Nucleoside antibiotics: biosynthesis, regulation, and biotechnology. *Trends Microbiol.* **2015**, *23*, 110.

Nix, D. E.; Swezey, R. R.; Hector, R.; Galgiani, J. N. Pharmacokinetics of nikkomycin Z after single rising oral doses. *Antimicrob Agents Chemother.* **2009**, *53*, 2517.

Noguchi, T.; Komoto, K.; Yasuda, Y.; Hashimoto, S.; Niida, T. New antibiotics, miharamycins A and B. III. Control activity of miharamycin against plant disease. *Meiji Seika Kenkyu Nenpo.* **1967**, *9*, 11.

Noguchi, T.; Yasuda, Y.; Niida, T.; Shomura, T. Inhibitory effects of Miharamycin A on the multiplication of plant viruses and the symptom development. *Annals of the Phytopathological Society of Japan.* **1968**, *34*, 323.

Oberthur, M.; Leimkuhler, C.; Kahne, D. A practical method for the stereoselective generation of beta-2-deoxy glycosyl phosphates. *Org Lett.* **2004**, *6*, 2873.

Oki, T.; Hirano, M.; Tomatsu, K.; Numata, K.; Kamei, H. Cispentacin, a new antifungal antibiotic. II. In vitro and in vivo antifungal activities. *J Antibiot (Tokyo).* **1989**, *42*, 1756.

Ortiz de Montellano, P. R.; Mathews, J. M.; Langry, K. C. Autocatalytic inactivation of cytochrome p-450 and chloroperoxidase by 1-aminobenzotriazole and other aryne precursors. *Tetrahedron* **1984**, *40*, 511

Otten, S. L.; Liu, X.; Ferguson, J.; Hutchinson, C. R. Cloning and characterization of the Streptomyces peucetius dnrQS genes encoding a daunosamine biosynthesis enzyme and a glycosyl transferase involved in daunorubicin biosynthesis. *J Bacteriol.* **1995**, *177*, 6688.

Otten, S. L.; Gallo, M. A.; Madduri, K.; Liu, X.; Hutchinson, C. R. Cloning and characterization of the Streptomyces peucetius dnmZUV genes encoding three enzymes required for biosynthesis of the daunorubicin precursor thymidine diphospho-L-daunosamine. *J Bacteriol.* **1997**, *179*, 4446.

Overbeek, R.; Olson, R.; Pusch, G. D.; Olsen, G. J.; Davis, J. J.; Disz, T.; Edwards, R. A.; Gerdes, S.; Parrello, B.; Shukla, M.; Vonstein, V.; Wattam, A. R.; Xia, F.; Stevens, R. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res.* **2014**, *42*, D206.

Peng, C.; Pu, J. Y.; Song, L. Q.; Jian, X. H.; Tang, M. C.; Tang, G. L. Hijacking a hydroxyethyl unit from a central metabolic ketose into a nonribosomal peptide assembly line. *Proc Natl Acad Sci U S A.* **2012**, *109*, 8540.

Peng, R.; VanNieuwenhze, M. S. A model study for constructing the DEF-benzoxocin ring system of menogaril and nogalamycin via a reductive Heck cyclization. *Org Lett.* *14*, 1962.

Perlman, D.; Vlietinck, A. J.; Matthews, H. W.; Lo, F. F. Microbial production of vitamin B12 antimetabolites. I. N5-hydroxy-L-arginine from Bacillus cereus 439. *J. Antibiot.* **1974**, *27*, 826.

The pET Expression System. Eleventh ed. 2006.

Powis, G. Free radical formation by antitumor quinones. *Free Radic Biol Med.* **1989**, *6*, 63.

Price, N. P.; Furukawa, T.; Cheng, F.; Qi, J.; Chen, W.; Crich, D. Biosynthesis of 4-aminoheptose 2-epimers, core structural components of the septacidins and spicamycins. *J Antibiot (Tokyo).* **2014**, *67*, 405.

The QIAexpressionist. Fifth ed. 2003.

QIAgen Miniprep Handbook. Fourth ed. 2015.

Quigley, G. J.; Wang, A. H.; Ughetto, G.; van der Marel, G.; van Boom, J. H.; Rich, A. Molecular structure of an anticancer drug-DNA complex: daunomycin plus d(CpGpTpApCpG). *Proc Natl Acad Sci U S A.* **1980**, *77*, 7204.

Raaijmakers, J. M.; Mazzola, M. Diversity and natural functions of antibiotics produced by beneficial and plant pathogenic bacteria. *Annu Rev Phytopathol.* **2012**, *50*, 403.

Rachakonda, S.; Cartee, L. Challenges in antimicrobial drug discovery and the potential of nucleoside antibiotics. *Curr Med Chem.* **2004**, *11*, 775.

Rangaswamy, V.; Jiralerspong, S.; Parry, R.; Bender, C. L. Biosynthesis of the Pseudomonas polyketide coronafacic acid requires monofunctional and multifunctional polyketide synthase proteins. *Proc Natl Acad Sci U S A.* **1998**, *95*, 15469.

Rangaswamy, V.; Mitchell, R.; Ullrich, M.; Bender, C. Analysis of genes involved in biosynthesis of coronafacic acid, the polyketide component of the phytotoxin coronatine. *J Bacteriol.* **1998**, *180*, 3330.

Robinson, H. J.; Graessle, O. E.; Smith, D. G. Studies on the Toxicity and Activity of Streptothricin. *Science.* **1944**, *99*, 540.

Rackham, E. J.; Gruschow, S.; Goss, R. J. Revealing the first uridyl peptide antibiotic biosynthetic gene cluster and probing pacidamycin biosynthesis. *Bioeng Bugs.* **2011**, *2*, 218.

Rackham, E. J.; Gruschow, S.; Ragab, A. E.; Dickens, S.; Goss, R. J. Pacidamycin biosynthesis: identification and heterologous expression of the first uridyl peptide antibiotic gene cluster. *Chembiochem.* **2010**, *11*, 1700.

Ragab, A. E.; Gruschow, S.; Tromans, D. R.; Goss, R. J. Biogenesis of the unique 4',5'-dehydronucleoside of the uridyl peptide antibiotic pacidamycin. *J Am Chem Soc.* **2011**, *133*, 15288.

Ramaley, R.; Fujita, Y.; Freese, E. Purification and properties of Bacillus subtilis inositol dehydrogenase. *J Biol Chem.* **1979**, *254*, 7684.

Reeves, A. R.; Brikun, I. A.; Cernota, W. H.; Leach, B. I.; Gonzalez, M. C.; Weber, J. M. Effects of methylmalonyl-CoA mutase gene knockouts on erythromycin production in carbohydrate-based and oil-based fermentations of Saccharopolyspora erythraea. *Journal of Industrial Microbiology and Biotechnology.* **2006**, *33*, 600.

Richards, S. It's more than stamp collecting: how genome sequencing can unify biological research. *Trends Genet.* **2015**, *31*, 411.

Rodriguez-Garcia, A.; Ludovice, M.; Martin, J. F.; Liras, P. Arginine boxes and the argR gene in Streptomyces clavuligerus: evidence for a clear regulation of the arginine pathway. *Mol Microbiol.* **1997**, *25*, 219.

Romo, A. J.; Liu, H. W. Mechanisms and structures of vitamin B(6)-dependent enzymes involved in deoxy sugar biosynthesis. *Biochim Biophys Acta.* **2011**, *1814*, 1534.

Ronen, A. 2-Aminopurine. *Mutat Res.* **1980**, *75*, 1.

Rutledge, P. J.; Challis, G. L. Discovery of microbial natural products by activation of silent biosynthetic gene clusters. *Nat Rev Microbiol.* **2015**, *13*, 509.

Sambrook, J.; Russell, D. W. *Molecular cloning : a laboratory manual*; Third ed.; Cold Spring Harbor Laboratory: Cold Spring Harbor, NY, 2001.

Samland, A. K.; Wang, M.; Sprenger, G. A. MJ0400 from Methanocaldococcus jannaschii exhibits fructose-1,6-bisphosphate aldolase activity. *FEMS Microbiol Lett.* **2008**, *281*, 36.

Saugar, I.; Sanz, E.; Rubio, M. A.; Espinosa, J. C.; Jimenez, A. Identification of a set of genes involved in the biosynthesis of the aminonucleoside moiety of antibiotic A201A from Streptomyces capreolus. *Eur J Biochem.* **2002**, *269*, 5527.

Schiemer, J.; Rev. A ed.; Tufts University Core Facility: 2014; Vol. 2014.

Schneider, S.; Sandalova, T.; Schneider, G.; Sprenger, G. A.; Samland, A. K. Replacement of a phenylalanine by a tyrosine in the active site confers fructose-6-phosphate aldolase activity to the transaldolase of Escherichia coli and human origin. *J Biol Chem.* **2008**, *283*, 30064.

Schneider, A.; Stachelhaus, T.; Marahiel, M. A. Targeted alteration of the substrate specificity of peptide synthetases by rational module swapping. *Mol Gen Genet.* **1998**, *257*, 308.

Sessa, C.; Gundersen, S.; ten Bokkel Huinink, W.; Renard, J.; Cavalli, F. Phase II study of intravenous menogaril in patients with advanced breast cancer. *J Natl Cancer Inst.* **1988**, *80*, 1066.

Seto, H.; Koyama, M.; Ogino, H.; Tsuruoka, T.; Inouye, S.; Otake, N. The structures of novel nucleoside antibiotics, miharamycin A and miharamycin B. *Tetrahedron.* **1983**, *24*, 1805.

Shen, B. Polyketide biosynthesis beyond the type I, II and III polyketide synthase paradigms. *Curr Opin Chem Biol.* **2003**, *7*, 285.

Shomura, T.; Hamamoto, K.; Ohashi, T.; Amano, S.; Yoshida, J.; Moriyama, C.; Niida, T. New antibiotics, miharamycins A and B. II. Some biological characteristics of miharamycin. *Meiji Seika Kenkyu Nenpo.* **1967**, *9*, 5.

Shubitz, L. F.; Roy, M. E.; Nix, D. E.; Galgiani, J. N. Efficacy of Nikkomycin Z for respiratory coccidioidomycosis in naturally infected dogs. *Med Mycol.* **2013**, *51*, 747.

Shubitz, L. F.; Trinh, H. T.; Perrill, R. H.; Thompson, C. M.; Hanan, N. J.; Galgiani, J. N.; Nix, D. E. Modeling nikkomycin Z dosing and pharmacology in murine pulmonary coccidioidomycosis preparatory to phase 2 clinical trials. *J Infect Dis.* **2014**, *209*, 1949.

Siitonen, V.; Blauenburg, B.; Kallio, P.; Mantsala, P.; Metsa-Ketela, M. Discovery of a two-component monooxygenase SnoaW/SnoaL2 involved in nogalamycin biosynthesis. *Chem Biol. 19,* 638.

Siitonen, V.; Claesson, M.; Patrikainen, P.; Aromaa, M.; Mantsala, P.; Schneider, G.; Metsa-Ketela, M. Identification of late-stage glycosylation steps in the biosynthetic pathway of the anthracycline nogalamycin. *Chembiochem.* **2012**, *13*, 120.

Silverman, R. *The Organic Chemistry of Enzyme-Catalyzed Reactions*; Revised ed.; Elsevier: Boston, 2002.

Simon, R.; Priefer, U.; Puhler, A. A broad host range mobilization system for in vivo genetic engineering: transposon mutagenesis in Gram negative bacteria. *Nature Biotechnology.* **1983**, *1*, 784.

Smith, B. Damage to the intrinsic cardiac neurones by rubidomycin (daunorubicin). *Br Heart J.* **1969**, *31*, 607.

Soltero, F. V.; Johnson, M. J. The effect of the carbohydrate nutrition on penicillin production by Penicillium chrysogenum Q-176. *Appl Microbiol.* **1953**, *1*, 52.

Sowers, L. C.; Boulard, Y.; Fazakerley, G. V. Multiple structures for the 2-aminopurine-cytosine mispair. *Biochemistry.* **2000**, *39*, 7613.

Stachelhaus, T.; Mootz, H. D.; Bergendahl, V.; Marahiel, M. A. Peptide bond formation in nonribosomal peptide biosynthesis. Catalytic role of the condensation domain. *J Biol Chem.* **1998**, *273*, 22773.

Stachelhaus, T.; Mootz, H. D.; Marahiel, M. A. The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. *Chem Biol.* **1999**, *6*, 493.

Stauffer, C. S.; Datta, A. Synthetic studies on amipurimycin: total synthesis of a thymine nucleoside analogue. *J Org Chem.* **2008**, *73*, 4166.

Stenland, C. J.; Lis, L. G.; Schendel, F. J.; Hahn, N. J.; Smart, M. A.; Miller, A. L.; von Keitz, M. G.; Gurvich, V. J. A practical and scalable manufacturing process for an anti-fungal agent, Nikkomycin Z. *Org Process Res Dev.* **2013**, *17*, 265.

Stirling, C. J. M. Leaving groups and nucleofugality in elimination and other organic reactions. *Acc Chem Res.* **1979**, *12*, 198.

Strieter, E. R.; Koglin, A.; Aron, Z. D.; Walsh, C. T. Cascade reactions during coronafacic acid biosynthesis: elongation, cyclization, and functionalization during Cfa7-catalyzed condensation. *J Am Chem Soc.* **2009**, *131*, 2113.

Sugimoto, Y.; Ishida, K.; Traitcheva, N.; Busch, B.; Dahse, H. M.; Hertweck, C. Freedom and constraint in engineered noncolinear polyketide assembly lines. *Chem Biol.* **2015**, *22*, 229.

Takahashi, A.; Ikeda, D.; Naganawa, H.; Okami, Y.; Umezawa, H. Bagougeramines A and B, new nucleoside antibiotics produced by a strain of Bacillus circulans. II.

Physico-chemical properties and structure determination. *J Antibiot (Tokyo).* **1986**, *39*, 1041.

Takahashi, H.; Liu, Y. N.; Chen, H.; Liu, H. W. Biosynthesis of TDP-l-mycarose: the specificity of a single enzyme governs the outcome of the pathway. *J Am Chem Soc.* **2005**, *127*, 9340.

Takahashi, H.; Liu, Y. N.; Liu, H. W. A two-stage one-pot enzymatic synthesis of TDP-L-mycarose from thymidine and glucose-1-phosphate. *J Am Chem Soc.* **2006**, *128*, 1432.

Takemoto, T.; Inamori, Y.; Kato, Y.; Kubo, M.; Morimoto, K.; Morisaka, K.; Sakai, M.; Sawada, Y.; Taniyama, H. Physiological activity of streptothricin antibiotics. *Chem Pharm Bull (Tokyo).* **1980**, *28*, 2884.

Tan, C.; Tasaka, H.; Yu, K. P.; Murphy, M. L.; Karnofsky, D. A. Daunomycin, an antitumor antibiotic, in the treatment of neoplastic disease. Clinical evaluation with special reference to childhood leukemia. *Cancer.* **1967**, *20*, 333.

Tello, M.; Rejzek, M.; Wilkinson, B.; Lawson, D. M.; Field, R. A. Tyl1a, a TDP-6-deoxy-D-xylo-4-hexulose 3,4-isomerase from Streptomyces fradiae: structure prediction, mutagenesis and solvent isotope incorporation experiments to investigate reaction mechanism. *Chembiochem.* **2008**, *9*, 1295.

Tercero, J. A.; Espinosa, J. C.; Lacalle, R. A.; Jimenez, A. The biosynthetic pathway of the aminonucleoside antibiotic puromycin, as deduced from the molecular analysis of the pur cluster of Streptomyces alboniger. *J Biol Chem.* **1996**, *271*, 1579.

Thibodeaux, C. J.; Melancon, C. E.; Liu, H. W. Unusual sugar biosynthesis and natural product glycodiversification. *Nature.* **2007**, *446*, 1008.

Thibodeaux, C. J.; Melancon, C. E., 3rd; Liu, H. W. Natural-product sugar biosynthesis and enzymatic glycodiversification. *Angew Chem Int Ed Engl.* **2008**, *47*, 9814.

Thoden, J. B.; Kappock, T. J.; Stubbe, J.; Holden, H. M. Three-dimensional structure of N5-carboxyaminoimidazole ribonucleotide synthetase: a member of the ATP grasp protein superfamily. *Biochemistry.* **1999**, *38*, 15480.

Thomas, G. J. In *Recent Progress in the Chemical Synthesis of Antibiotics*; First ed.; Lukacs, G., Ohno, M., Ed.; Springer-Verlag: Berlin, 1990, p 467.

Tokunaga, H.; Kojima, M.; Kuroha, T.; Ishida, T.; Sugimoto, K.; Kiba, T.; Sakakibara, H. Arabidopsis lonely guy (LOG) multiple mutants reveal a central role of the LOG-dependent pathway in cytokinin activation. *Plant J.* **2012**, *69*, 355.

Torkkell, S.; Kunnari, T.; Palmu, K.; Mantsala, P.; Hakala, J.; Ylihonko, K. The entire nogalamycin biosynthetic gene cluster of Streptomyces nogalater: characterization of a 20-kb DNA region and generation of hybrid structures. *Mol Genet Genomics.* **2001**, *266*, 276.

Torkkell, S.; Ylihonko, K.; Hakala, J.; Skurnik, M.; Mantsala, P. Characterization of Streptomyces nogalater genes encoding enzymes involved in glycosylation steps in nogalamycin biosynthesis. *Mol Gen Genet.* **1997**, *256*, 203.

Treede, I.; Hauser, G.; Muhlenweg, A.; Hofmann, C.; Schmidt, M.; Weitnauer, G.; Glaser, S.; Bechthold, A. Genes involved in formation and attachment of a two-carbon chain as a component of eurekanate, a branched-chain sugar moiety of avilamycin A. *Appl Environ Microbiol.* **2005**, *71*, 400.

Tsuruoka, T.; Yumoto, H.; Ezaki, N.; Niida, T. New antibiotics, miharamycins A and B. I. Isolation and characterization of miharamycins A and B. *Meiji Seika Kenkyu Nenpo.* **1967**, *9*, 1.

van Berkel, W. J.; Kamerbeek, N. M.; Fraaije, M. W. Flavoprotein monooxygenases, a diverse class of oxidative biocatalysts. *J Biotechnol.* **2006**, *124*, 670.

Vejpongsa, P.; Yeh, E. T. Prevention of anthracycline-induced cardiotoxicity: challenges and opportunities. *J Am Coll Cardiol.* **2014**, *64*, 938.

Vejpongsa, P.; Yeh, E. T. Topoisomerase 2beta: a promising molecular target for primary prevention of anthracycline-induced cardiotoxicity. *Clin Pharmacol Ther.* **2014**, *95*, 45.

Waksman, S. A. Production and Activity of Streptothricin. *J Bacteriol.* **1943**, *46*, 299.

Walsh, C. T. Insights into the chemical logic and enzymatic machinery of NRPS assembly lines. *Nat Prod Rep.* **2015**.

Walsh, C. T.; Chen, H.; Keating, T. A.; Hubbard, B. K.; Losey, H. C.; Luo, L.; Marshall, C. G.; Miller, D. A.; Patel, H. M. Tailoring enzymes that modify nonribosomal peptides during and after chain elongation on NRPS assembly lines. *Curr Opin Chem Biol.* **2001**, *5*, 525.

Walsh, C. T.; Fischbach, M. A. Natural products version 2.0: connecting genes to molecules. *J Am Chem Soc.* **2010**, *132*, 2469.

Walsh, C. T.; Zhang, W. Chemical logic and enzymatic machinery for biological assembly of peptidyl nucleoside antibiotics. *ACS Chem Biol.* **2011**, *6*, 1000.

Wang, J.; Zhang, X.; Ma, D.; Lee, W. N.; Xiao, J.; Zhao, Y.; Go, V. L.; Wang, Q.; Yen, Y.; Recker, R.; Xiao, G. G. Inhibition of transketolase by oxythiamine altered dynamics of protein signals in pancreatic cancer cells. *Exp Hematol Oncol.* **2013**, *2*, 18.

Weber, T.; Blin, K.; Duddela, S.; Krug, D.; Kim, H. U.; Bruccoleri, R.; Lee, S. Y.; Fischbach, M. A.; Muller, R.; Wohlleben, W.; Breitling, R.; Takano, E.; Medema, M. H. antiSMASH 3.0-a comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids Res.* **2015**, *43*, W237.

White-Phillip, J.; Thibodeaux, C. J.; Liu, H. W. Enzymatic synthesis of TDP-deoxysugars. *Methods Enzymol.* **2009**, *459*, 521.

Williams, L. D.; Egli, M.; Qi, G.; Bash, P.; van der Marel, G. A.; van Boom, J. H.; Rich, A.; Frederick, C. A. Structure of nogalamycin bound to a DNA hexamer. *Proc Natl Acad Sci U S A.* **1990**, *87*, 2225.

Winn, M.; Goss, R. J.; Kimura, K.; Bugg, T. D. Antimicrobial nucleoside antibiotics targeting cell wall assembly: recent advances in structure-function studies and nucleoside biosynthesis. *Nat Prod Rep.* **2010**, *27*, 279.

Wu, J.; Li, L.; Deng, Z.; Zabriskie, T. M.; He, X. Analysis of the mildiomycin biosynthesis gene cluster in Streptoverticillum remofaciens ZJU5119 and characterization of MilC, a hydroxymethyl cytosyl-glucuronic acid synthase. *Chembiochem.* **2012**, *13*, 1613.

Yoon, Y. J.; Beck, B. J.; Kim, B. S.; Kang, H. Y.; Reynolds, K. A.; Sherman, D. H. Generation of multiple bioactive macrolides by hybrid modular polyketide synthases in Streptomyces venezuelae. *Chem Biol.* **2002**, *9*, 203.

Zarins-Tutt, J. S.; Barberi, T. T.; Gao, H.; Mearns-Spragg, A.; Zhang, L.; Newman, D. J.; Goss, R. J. Prospecting for new bacterial metabolites: a glossary of approaches for inducing, activating and upregulating the biosynthesis of bacterial cryptic or silent natural products. *Nat Prod Rep.* **2015**.

Zeng, Y.; Kulkarni, A.; Yang, Z.; Patil, P. B.; Zhou, W.; Chi, X.; Van Lanen, S.; Chen, S. Biosynthesis of albomycin delta(2) provides a template for assembling siderophore and aminoacyl-tRNA synthetase inhibitor conjugates. *ACS Chem Biol.* **2012**, *7*, 1565.

Zeng, Y.; Roy, H.; Patil, P. B.; Ibba, M.; Chen, S. Characterization of two seryl-tRNA synthetases in albomycin-producing Streptomyces sp. strain ATCC 700974. *Antimicrob Agents Chemother.* **2009**, *53*, 4619.

Zerbino, D. R. Using the Velvet de novo assembler for short-read sequencing technologies. *Curr Protoc Bioinformatics.* **2010**, *Chapter 11*, Unit 11 5.

Zerbino, D. R.; Birney, E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* **2008**, *18*, 821.

Zerbino, D. R.; McEwen, G. K.; Margulies, E. H.; Birney, E. Pebble and rock band: heuristic resolution of repeats and scaffolding in the velvet short-read de novo assembler. *PLoS One.* **2009**, *4*, e8407.

Zhang, G.; Zhang, H.; Li, S.; Xiao, J.; Zhu, Y.; Niu, S.; Ju, J.; Zhang, C. Characterization of the amicetin biosynthesis gene cluster from Streptomyces vinaceusdrappus NRRL 2363 implicates two alternative strategies for amide bond formation. *Appl Environ Microbiol.* **2012**, *78*, 2393.

Zhang, Q.; Gould, S. J.; Zabriskie, T. M. A new cytosine glycoside from Streptomyces griseochromogenes produced by the use in vivo of enzyme inhibitors. *J Nat Prod.* **1998**, *61*, 648.

Zhang, W.; Ntai, I.; Bolla, M. L.; Malcolmson, S. J.; Kahne, D.; Kelleher, N. L.; Walsh, C. T. Nine enzymes are required for assembly of the pacidamycin group of peptidyl nucleoside antibiotics. *J Am Chem Soc.* **2011**, *133*, 5240.

Zhang, W.; Ntai, I.; Kelleher, N. L.; Walsh, C. T. tRNA-dependent peptide bond formation by the transferase PacB in biosynthesis of the pacidamycin group of pentapeptidyl nucleoside antibiotics. *Proc Natl Acad Sci U S A.* **2011**, *108*, 12249.

Zhang, W.; Ostash, B.; Walsh, C. T. Identification of the biosynthetic gene cluster for the pacidamycin group of peptidyl nucleoside antibiotics. *Proc Natl Acad Sci U S A.* **2010**, *107*, 16828.

Zhang, Y.; Morar, M.; Ealick, S. E. Structural biology of the purine biosynthetic pathway. *Cell Mol Life Sci.* **2008**, *65*, 3699.

Zhao, Q.; He, Q.; Ding, W.; Tang, M.; Kang, Q.; Yu, Y.; Deng, W.; Zhang, Q.; Fang, J.; Tang, G.; Liu, W. Characterization of the azinomycin B biosynthetic gene cluster revealing a different iterative type I polyketide synthase for naphthoate biosynthesis. *Chem Biol.* **2008**, *15*, 693.

Zhu, Q.; Li, J.; Ma, J.; Luo, M.; Wang, B.; Huang, H.; Tian, X.; Li, W.; Zhang, S.; Zhang, C.; Ju, J. Discovery and engineered overproduction of antimicrobial nucleoside antibiotic A201A from the deep-sea marine actinomycete Marinactinospora thermotolerans SCSIO 00652. *Antimicrob Agents Chemother.* **2012**, *56*, 110.

# Vita

In 2009 Anthony James Romo graduated from the College of Pharmacy of the University of Texas at Austin with a Pharm.D. During his pharmacy school coursework, Anthony worked in the laboratory of Dr. Hung-wen (Ben) Liu, where he gained an appreciation for enzymology and the biosynthesis of unusual sugar-containing natural products. Upon graduation from pharmacy school, Anthony was admitted to the pharmaceutical sciences graduate program in the College of Pharamcy at the University of Texas at Austin. Largely as a result of his experience during pharmacy school, Anthony joined the research group of Dr. Ben Liu in the fall of 2009. In the ensuing six years, Anthony worked to uncover the biosynthesis of unusual peptidyl nucleoside antibiotics. He was the recipient of a Graduate School Diversity Fellowship and the Josephina Lesvia Falcon Graduate Fellowship.

Permanent email address: anthonyromo@utexas.edu

This dissertation was typed by Anthony James Romo.